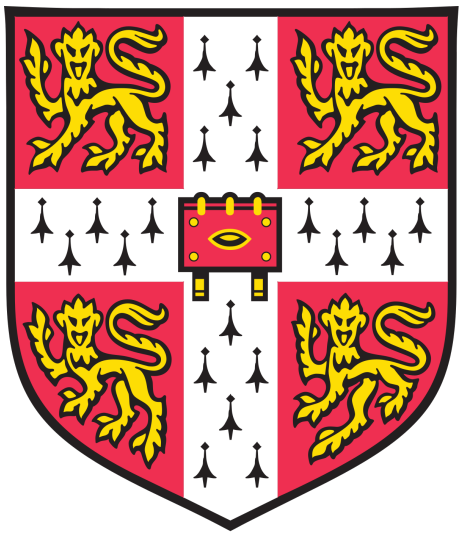
A thesis submitted for the degree of Doctor of Philosophy

Genetically Dissecting Disease Interactions between *Parastagonospora nodorum* and Wheat (*Triticum aestivum* L.)



Rowena Cathryn Downie

Magdalene College
Cambridge

University of Cambridge

Supervised by
Dr James Cockram, NIAB
Prof. Richard Oliver, CCDM, Curtin University
Dr Ian Henderson, Department of Plant Sciences, University of Cambridge

12th September 2019

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. The text does not exceed 60,000 words in length.

Rowena C. Downie

Abstract

Genetically dissecting disease interactions between *Parastagonospora nodorum* and wheat (*Triticum aestivum* L.) - Rowena C. Downie

Parastagonospora nodorum is a necrotrophic fungal pathogen of wheat (*Triticum aestivum* L.) and the causative agent of stagonospora nodorum blotch (SNB). *P. nodorum* mediates host cell death using proteinaceous effectors. Three *P. nodorum* effectors have been cloned: SnToxA, SnTox1 and SnTox3. SnTox3 susceptible wheat cultivars possess a dominant susceptibility allele at the *Snn3-B1* locus; breeding to remove SnTox3 sensitivity could help increase *P. nodorum* disease resistance. The research presented here falls into three main areas:

1. Genetic mapping of SnTox3 sensitivity: the major SnTox3 sensitivity locus *Snn3-B1* was fine-mapped using an association mapping panel, the Avalon × Cadenza population and a MAGIC population. These defined a 6.2 kb region on chromosome 5B explaining ≥ 32 % of the phenotypic variation. The most significantly associated SNP was converted to a KASP marker. Additionally, nine minor SnTox3 sensitivity QTL were identified on chromosomes 1B, 2A, 2B, 3B, 4D, 6A, 6B and 7B, each accounting for 2.6–6.0 % of the phenotypic variation.
2. Gene expression analysis: combining RNA-seq and genomic analysis in a time-course experiment using a SnTox3 sensitive (Cadenza) and insensitive variety (Avalon), two treatments (+/- SnTox3 infiltration) and 0, 4, 8, 12 and 24 hour time-points allowed investigation of gene expression before and after SnTox3 infiltration. Two *Snn3-B1* candidate genes were identified, a protein kinase and wall-associated kinase, with analysis of genome-wide patterns of gene expression providing insights into the molecular mechanisms of the SnTox3-*Snn3-B1* pathway.
3. Adult plant resistance: The Avalon × Cadenza bi-parental population was used to investigate adult plant *P. nodorum* resistance. First, an improved Avalon × Cadenza genetic map was developed using publicly available genotypic data. This was combined with field trial data to identify a major QTL on chromosome 5A for leaf and glume blotch and a further seven QTL on chromosomes 1B, 2D, 3B, 5A, 5B and 7B. Of these, field resistance QTL were found to co-locate with the effector sensitivity QTL *Snn1* and *Snn7*.

Collectively, this research provides wheat breeders and researchers with knowledge and molecular resources regarding sensitivity to the *P. nodorum* effector, SnTox3, as well as field QTL for both leaf and glume blotch.

Acknowledgements

THANK you to my supervisors: James Cockram, Richard Oliver, and Ian Henderson for all their expertise, help, and guidance throughout my PhD. I also owe my thanks to my colleagues and collaborators, including my advisors Kar-Chun Tan and Huyen Phan, as well as Lawrence Percival-Alwyn, for the mentoring and guidance they provided over the last few years. Additionally, I am grateful to my collaborators at NMBU, particularly Morten Lillemo and the staff at Vollebekk who oversaw my field trials. I am thankful for my funding from the BBSRC DTP as well as additional funding from CCDM and Magdalene College and for the support from NIAB, the University of Cambridge and Magdalene College.

I am also incredibly thankful to the former PhD office and middle office for the camaraderie and encouragement that always brightened my day. A huge thank you to the MCR of Magdalene College whom I will never forget as well as Magdalene Boat Club, both gave me a huge amount of mental strength to keep going. For all the encouragement and support, I am hugely grateful to Ellie. Thank you to my family, Mum, Dad, Lynsey, Harro, Dagmar, Henning, Christiane, Steffi, Marcus, Benno, Mila, Anni and Bailey for always being there and providing wonderful distractions.

Finally, but most importantly, I owe so much to my rock, who has amazingly helped me to stay level-headed, even in the final weeks of writing. So, I dedicate this to you, mein Verlobter, Oliver.

This thesis was typeset in LaTeX.

Contents

List of Figures

List of Tables

List of Abbreviations

aa	amino acid
AM	association mapping
bp	base pair
CDS	coding sequence
CIM	composite interval mapping
DArT	diversity arrays technology
DEG	differentially expressed gene
EGF	epidermal growth factor
EGF_CA	calcium binding epidermal growth factor
ETI	effector-triggered immunity
ETS	effector-triggered susceptibility
GO	gene ontology
GS	growth stage
GUB	galacturonic acid binding
GWAS	genome-wide association study
hpi	hours post-infiltration
IM	interval mapping
KASP	Kompetitive Allele Specific PCR
LOD	logarithm of the odds
LRR	leucine-rich repeat
MAGIC	multi-parent advanced generation inter-cross
Mlo	Mildew Locus O
NBS	nucleotide binding site
NIL	near isogenic line
PAMP	pathogen associated molecular pattern
PCA	principal component analysis

PCR	polymerase chain reaction
PKC	protein kinase C
PRR	pattern recognition receptor
PTI	PAMP-triggered immunity
PVE	percentage variation explained
QTL	quantitative trait locus
R	resistance
RIL	recombinant inbred lines
ROS	reactive oxygen species
S/TPK	serine/threonine protein kinase
SMA	single marker analysis
SNB	Stagonospora Nodorum Blotch
SNP	single nucleotide polymorphism
SSR	simple sequence repeat
STKc-IRAK	serine/threonine kinase - interleukin-1 receptor-associated kinase
TILLING	Targeting Induced Local Lesions In Genomes
UTR	untranslated region
WAK	wall associated kinase

1

Introduction

Bread wheat, *Triticum aestivum* L. (2n = 6x = 42, AABBDD), is a crop of great global value, with a production of 728 million tonnes per annum (Li et al., 2007; Cui et al., 2014; FAO, 2015). It is grown across most of the world, including temperate regions up to latitudes of 67° N and regions of the southern hemisphere with latitudes as southerly as 45° S (Shewry, 2009). It is vital to food security that production of staple crops, such as wheat, are able to meet the demands of a rapidly growing human population, currently predicted to reach 9.6 billion by 2050 (UN, 2013). This will require an estimated increase in cereal production of one billion tonnes per annum (FAO, 2009). Numerous complications hinder production increases, including legislation restricting certain fungicide and pesticide usage as well as reductions in arable land. As a result, there has never been a greater need for increases in food production per hectare, whilst minimising inputs: a process known as sustainable intensification (Araus et al., 2008; UN, 2013). Improving the genetic resistance of wheat to the pathogens responsible for global yield losses is an important component of sustainable intensification approaches.

1.1 Wheat Evolution and Genetics

This common hexaploid crop, bread wheat, has come into being as a result of two hybridisation events, contributing to its polyploidy status (IWGSC et al., 2014). It is surmised that the initial polyploidisation event occurred, around 0.82 million

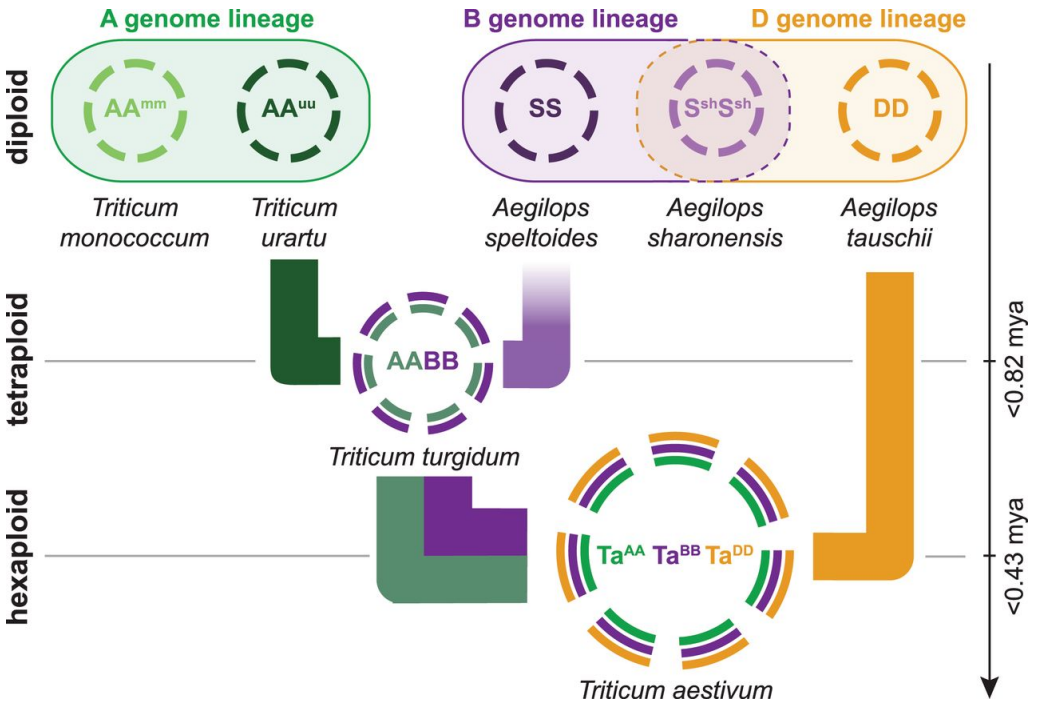


Figure 1.1 – The evolution of *Triticum aestivum*. Showing two hybridisation events resulting in the final allohexaploid bread wheat. Taken from IWGSC et al., 2014.

years ago, between the A sub-genome donor an ancestral line of *Triticum urartu* and the B sub-genome donor an ancestral line of *Aegilops speltoides*. Followed by a second polyploidisation event between the resulting tetraploid AABB *Triticum turgidum* and *Aegilops tauschii*, the D sub-genome donor, around 0.43 million years ago (IWGSC et al., 2014; Loginova and Silkova, 2018; Figure 1.1). However many researchers still say that the B sub-genome donor is unknown; the sequences of the B sub-genome greater diversification rate is a possible contributing factor to the increased difficulty in B genome progenitor identification (Petersen et al., 2006; Loginova and Silkova, 2018). These polyploidisation events therefore contribute to a very large overall genome size of approximately 17 Gbp (Krasileva et al., 2013). Within which, wheat homoeologues share on average 95 % sequence identity, making it a challenge to distinguish between the sub-genomes (Borrill et al., 2015). Although this large complex allohexaploid genome structure has been merited as key in wheat's ability to adapt to a broad range of climatic conditions (IWGSC et al., 2014).

1.2 Fungal Pathogens of Wheat

Within an agricultural setting, wheat is exposed to a whole plethora of pathogens, ranging from bacterial and viral diseases to parasitic nematodes and fungi. However, the fungal pathogens, which will be the focus here, fall into three distinct categories:

biotrophic, hemibiotrophic and necrotrophic (Table 1.1; Tang et al., 2018). Biotrophs are specialist obligate pathogens that require a living host to feed, grow and reproduce (Lorrain et al., 2019). Although whilst host death may be avoided, crop growth is often stunted (Hane et al., 2020). These fungal pathogens are distinctive due to their haustoria, following entry into the host the biotroph starts to form the specialised haustorial structure that will allow for nutrient uptake from the plant (Lorrain et al., 2019). Two key examples of destructive biotrophs that infect wheat are *Blumeria graminis* (powdery mildew) and the *Puccinia* species (the rusts) (Tang et al., 2018).

Hemibiotrophic pathogens display characteristics of a typical biotroph, but also characteristics of necrotrophs, able to move sequentially from an asymptomatic biotrophic phase to a destructive necrotrophic phase (Chowdhury et al., 2017). This biotrophy-necrotrophy switch allows the pathogen to be facultative in nature, feeding initially on living tissue and later on dead tissue, through the use of its haustorium or haustorium-like structure (Hane et al., 2020). Key examples of hemibiotrophic pathogens of wheat are *Zymoseptoria tritici* (septoria) and the *Fusarium* species (Duba et al., 2018).

Finally the third category of fungal pathogens, the necrotrophs, are contrastingly facultative pathogens that feed without the need for haustoria, but exclusively on dying host cells and tissues (Hane et al., 2020). Key examples of necrotrophic fungal pathogens in wheat are *Pyrenophora tritici-repentis* (tan spot) and *Parastagonospora nodorum* (stagonospora nodorum blotch). The necrotrophs differ in particular with regard to the host-specific toxins and necrotrophic effectors they produce to trigger host cell death, where as biotrophs (and hemibiotrophs) typically produce avirulence effectors that stimulate a qualitative response in the host (Duba et al., 2018). However in terms of host resistance, a large part of plant defence against biotrophs is programmed cell death (PCD) and in turn the up-regulation of the salicylic acid-dependant pathway which is known to regulate defence responses (Oliver and Ipcho, 2004). Whereas necrotrophs benefit from host cell death, and so PCD is of no benefit to host in terms of defence against a necrotrophic pathogen. But rather by activating jasmonic acid signalling pathways in a quantitative response to upregulate to defence (Hane et al., 2020). However, the fungal pathogen's attack is to secret effectors or toxins to alter the host's metabolism in order to evade defence mechanisms or to manipulate plant gene transcription to benefit to the pathogen (Möller and Stukenbrock, 2017).

1.3 Stagonospora Nodorum Blotch

Stagonospora Nodorum Blotch (SNB) is a cereal disease of global economic importance, and is caused by the necrotrophic fungal pathogen *Parastagonospora*

Table 1.1 – Fungal pathogen traits

Trait	Biotroph	Hemibiotroph	Necrotroph
Lifestyle	Obligate	Facultative	Facultative
Nutrient source	Living host cells	Living and dead host cells	Dead host cells
Feeding structures	Haustoria	Haustoria(-like) structures	No haustoria
Effectors	Avirulence effectors	Avirulence effectors	Host-specific toxins nectrophic effectors
Host defence hormones	Salicylic acid	Salicylic and Jasmonic acid	Jasmonic acid
Host resistance genes	Qualitative	Qualitative	Quantitative

A summary of the key characteristics of the three fungal pathogen types (Hane et al., 2020; Lorrain et al., 2019; Duba et al., 2018).

nodorum (syn.: *Phaeosphaeria nodorum*) (Quaedvlieg et al., 2013). The visual symptoms of SNB are chlorosis and necrosis of the leaf tissue, as well as discoloration and necrosis of the glumes, often in the form of lesions (Solomon et al., 2006). These symptoms reduce the leaf surface area capable of photosynthesis, therefore limiting overall crop growth and yield. SNB is known to cause yield losses of up to 31 % (Bhathal et al., 2003). In recent decades, there has been a focal shift in UK prevalence from *P. nodorum* to another necrotrophic wheat pathogen, *Zymoseptoria tritici* (syn.: *Septoria tritici*, *Mycosphaerella graminicola*). The underlying reasons for this change are not fully understood, but have been suggested to be due to increased levels of *Z. tritici* host susceptibility, changes in climate, increased fertiliser use and SO_2 emissions (West et al., 2012; Shaw et al., 2008). However, *P. nodorum* remains an important vector of disease worldwide, with current economic losses to wheat production being most marked in Scandinavia, the USA and Australia (Eyal, 1999; Hane et al., 2007; Murray and Brennan, 2009). Furthermore, it appears that *P. nodorum* is still moving into new niches, for example, in 2017 it was observed for the first time on emmer wheat (*Triticum dicoccoides*) in Turkey, and due to changing climatic conditions, SNB has now become a major problem in Himachal Pradesh, India (Cat et al., 2018, Katoch et al., 2018). In addition, *P. nodorum* — along with the necrotrophic pathogen *Pyrenophora tritici-repentis* — is the first of the Dothiodeomycete class of fungal pathogens to have had their genomes sequenced, and so may serve as a model for the necrotrophic pathogen lifecycle (Liu et al., 2009).

1.4 *Parastagonospora nodorum* Lifecycle and Infection Process

P. nodorum is a necrotrophic fungal pathogen that assimilates nutrients released after host cell death (De Wit et al., 2009). As a necrotrophic pathogen, *P. nodorum* produces proteinaceous effectors to fuel disease progression and triggering the plant's

receptors to promote sensitivity and tissue death (De Wit et al., 2009). *P. nodorum* has both asexual and sexual cycles (Figure 1.2). As part of the asexual cycle, fruiting bodies, called pycnidia, form in lesions on the leaf to promote spore development for local dispersal. In contrast, the sexual life cycle produces ascospores, derived from pseudothecium, that allow long distance aerial dispersion. The presence of both sexual and asexual reproduction mechanisms is hypothesised to provide *P. nodorum* with a high evolutionary potential, which could account for the many effectors it produces; an alternative theory is that effectors were acquired via horizontal gene transfer (McDonald and Linde, 2002; McDonald et al., 2013). Semi-dwarf varieties of wheat may have a higher risk of secondary *P. nodorum* infection due to the close vertical spacing of the leaves, as conidia, produced by pycnidia, are sent on an upward trajectory by water droplets (Bahat et al., 1980; Eyal, 1981; Brennan et al., 1985). This is particularly relevant as the majority of modern wheat varieties have been bred for their short (or semi-dwarf) stature.

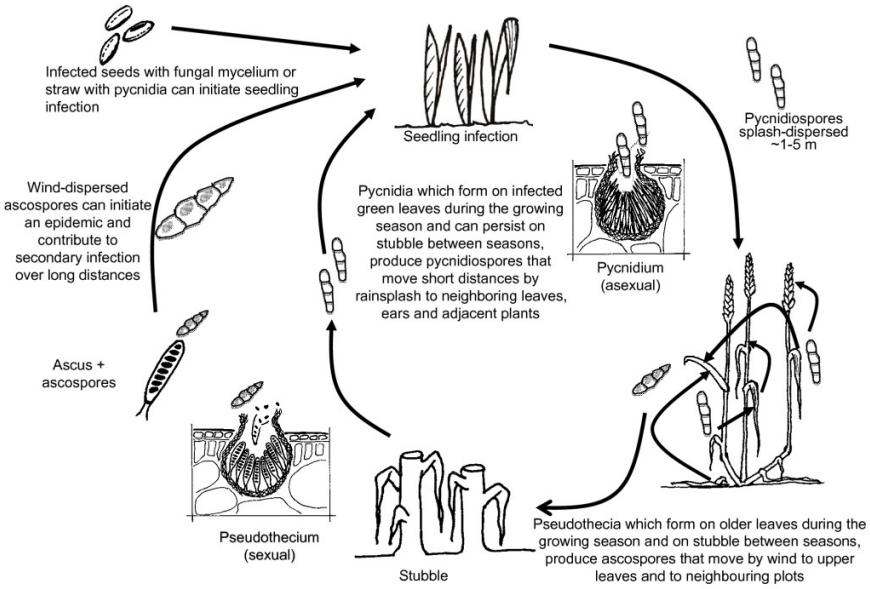


Figure 1.2 – The life cycle of *Parastagonospora nodorum*. Showing sexual and asexual cycles that produce ascospores and pycnidiospores respectively. Taken from Sommerhalder et al., 2011.

1.5 Stagonospora Nodorum Blotch Disease Management

Currently SNB is controlled, at least in part, by fungicides. It has been found that in years with greater rainfall in early spring, *P. nodorum* infected fields treated with the fungicides propiconazole and tebuconazole had greater yield responses (Salam et al., 2013). However, it has since been reported that two non-synonymous mutations in the *P. nodorum* gene *CYP51* are correlated with significantly reduced sensitivity to propiconazole (Pereira et al., 2017). An alternative class of fungicide for SNB disease management are strobilurins (e.g. azoxystrobin) (Eyal, 1999). Nevertheless, strobilurin resistance has been found in many other wheat fungal pathogens, likely due to strong selection pressure as strobilurins inhibit mitochrondrial respiration (Solomon et al., 2006). Furthermore, it has been reported that reduced sensitivity to strobilurins has been found in some *P. nodorum* isolates, although this observation was not upheld *in vitro* (Blixt et al., 2009). This reduced sensitivity is thought to be linked to the amino acid substitution G143A in the *P. nodorum* gene *cytochrome b*, as this mutation dominated fields that had been regularly treated with strobilurins (Blixt et al., 2009). However, it has been found that azoxystrobin combined with the chemosensitising agent, thymol (2-isopropyl-5-methylphenol) contributed to greater levels of *P. nodorum* growth inhibition and so could be used when faced with an azoxystrobin resistant strain (Dzhavakhiya et al., 2012). As the infection process becomes better understood, this offers more opportunities to also explore host genetic resistance. At present, however, chemical control still remains a critical part of SNB disease management.

1.6 Fungal Effector Proteins and Sensitivity Loci

As a necrotrophic fungus, *P. nodorum* lives off dead plant tissue, using fungal effector proteins (also known as 'host selective toxins') to produce a hypersensitive response in the form of programmed cell death. This is genetically induced using an 'inverse gene for gene system' through effector-triggered susceptibility (ETS) (Friesen et al., 2007; Oliver et al., 2012; Liu et al., 2009). Although inverse gene for gene systems are uncommon, they are certainly not unheard of amongst the necrotrophs, as they are also found in *P. tritici-repentis*, *Pyrenophora teres* and *Bipolaris sorokiniana*, the causal agents of tan spot in wheat, net blotch in barley and spot blotch in wheat and barley respectively (Richards et al., 2017). Plant-pathogen interactions are usually quantitative in nature, and therefore breeding for disease resistant cultivars is typically a long and difficult process. However, if pathogen effectors can be identified and functionally characterised, important components of host-pathogen interactions

could be broken-down into their constituent parts. This would allow targeted breeding to eliminate individual wheat effector sensitivity loci in order to produce resistant cultivars in areas of high disease prevalence, including areas with fungicide resistant *P. nodorum* strains. Effector proteins were first described in 1933 through the study of the host-pathogen interaction between *Alternaria alternata* and the Japanese pear, *Pirus serotine* (Tanaka, 1933). Evidence of eight effectors has been found in *P. nodorum*, these are designated SnToxA, SnTox1, SnTox2, SnTox3, SnTox4, SnTox5, SnTox6 and SnTox7, along with corresponding *T. aestivum* sensitivity loci, *Tsn1*, *Snn1*, *Snn2*, *Snn3*, *Snn4*, *Snn5*, *Snn6* and *Snn7*, respectively (Friesen et al., 2012; Gao et al., 2015; Shi et al., 2015). Of these effectors, only three have been identified at the gene level. SnTox1 was the first *P. nodorum* effector to be identified. It was isolated from culture filtrate of the aggressive isolate, Sn2000, collected from wheat fields in North Dakota, USA. SnTox1 was identified as a protein due to its inactivity following exposure to proteinase K, which cleaves peptide bonds. The wheat sensitivity locus for SnTox1 was mapped to the short arm of chromosome 1B and designated *Snn1* using a bi-parental mapping population with the synthetic wheat, W-7984, and the spring wheat, Opata 85, (sensitive and insensitive to SnTox1, respectively) as parental lines (Liu et al., 2004). The *Snn1* locus has recently been fine-mapped using a genetically diverse wheat multi-parent advanced generation intercross (MAGIC) population to the short arm of chromosome 1B with an additional quantitative trait locus (QTL) identified on the long arm of chromosome 5A (Mackay et al., 2014; Cockram et al., 2015). This allowed Kompetitive Allele Specific PCR (KASP) markers to be developed that can easily be used by breeders to predict the allelic state at *Snn1*. The Sn2000 isolate was also shown to produce SnToxA, another effector protein, which is similar in structure and function to PtrToxA, an effector produced by *P. tritici-repentis*. Sensitivity to SnToxA was mapped as a major QTL to chromosome 5B and the locus designated *Tsn1* (Liu et al., 2006). The most recently identified effector gene was *SnTox3*, which was found to have a corresponding host sensitivity locus on chromosome 5B, termed *Snn3-B1* (Friesen et al., 2008).

1.6.1 SnToxA-*Tsn1* Interaction

SnToxA was first discovered to be secreted by *P. nodorum* in 2006, and *SnToxA* was found to have 99.7 % DNA sequence similarity to *PtrToxA* of *P. tritici-repentis*. Interestingly, due to the monomorphism of *PtrToxA* compared to the high level of diversity of *SnToxA*, it is likely that *ToxA* was introduced into the *P. tritici-repentis* genome through interspecific gene transfer from *P. nodorum* (Friesen et al., 2006). The corresponding host sensitivity locus for SnToxA, *Tsn1*, was first discovered

in 1996 as conferring sensitivity to the *P. tritici-repentis* PtrToxA effector. *Tsn1* was later confirmed as the corresponding host sensitivity locus for SnToxA after the identification and characterisation of SnToxA (Faris et al., 1996, Liu et al., 2006). The recent cloning of *Tsn1* has found the genomic region of 10,581 base pairs (bp) to encode a protein of 1490 amino acids (aa) with three predicted domains: a serine/threonine protein kinase (S/TPK) (with ATP binding, substrate binding site and activation loop), a nucleotide binding site (NBS) and 24 leucine-rich repeats (LRRs) located across eight exons (Figure 1.3; Faris et al., 2010). While NBS-LRRs form the largest class of plant resistance (R) genes, unlike most monocots' R genes, *Tsn1* does not encode a coiled-coil domain (Faris et al., 2010). SnToxA was found to interact with the dimeric PR-1-type pathogenesis-related protein, TaPR-1-5, to activate *Tsn1*-controlled cell death pathways (Breen et al., 2016). *Tsn1* expression has been found to be regulated by light and the circadian clock, which could also provide an explanation for the light dependent nature of the SnToxA-*Tsn1* interaction (Faris et al., 2010, Manning and Ciuffetti, 2005). Interestingly, ToxA has also been discovered in *B. sorokiniana*, which raises the question whether *Tsn1* is also found in barley (Friesen et al., 2018).

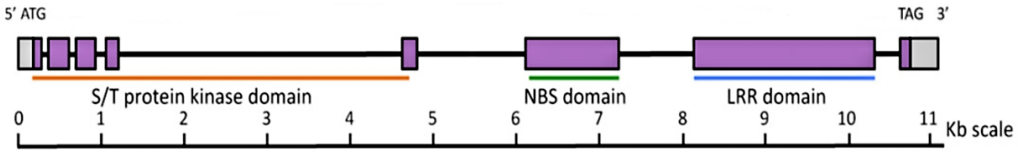


Figure 1.3 – The structure of *Tsn1*. Showing exons in purple with the regions spanning the exons encoding the serine/threonine (S/T) protein kinase underlined in orange, a nucleotide binding site (NBS) in green and a leucine rich repeat (LRR) domain in blue. Untranslated regions are shown in grey. Adapted from Faris et al., 2010.

1.6.2 SnTox1-*Snn1* Interaction

The *P. nodorum* effector SnTox1 is 117 aa in length with an initial 17 aa signal sequence and a high cysteine residue content with 16 cysteine residues in total (Liu et al., 2012). After cleavage a mature 10.33 kDa protein remains (Liu et al., 2012). SnTox1 and the SnTox1-*Snn1* interaction were first reported in 2004, after partially purifying a novel protein effector and discovering the SnTox1 sensitivity QTL, *Snn1*, on the short arm of chromosome 1B (Liu et al., 2004). *Snn1* was later map-based cloned and found to encode a galacturonic acid binding (GUB) wall associated kinase (WAK), and to possess calcium binding epidermal growth factor (EGF_CA) and S/TPK domains (Figure 1.4; Shi et al., 2016b). S/TPKs form another (smaller) class of plant R proteins (Faris et al., 2010). Interestingly, two non-synonymous mutations

in the SnTox1 insensitive *snn1* allele were found at amino acid residues 347 and 429 in the predicted protein (Shi et al., 2016b). Unlike the case for the SnToxA-*Tsn1* interaction, SnTox1 is thought to interact directly with Snn1 (Figure 1.5; Shi et al., 2016b).

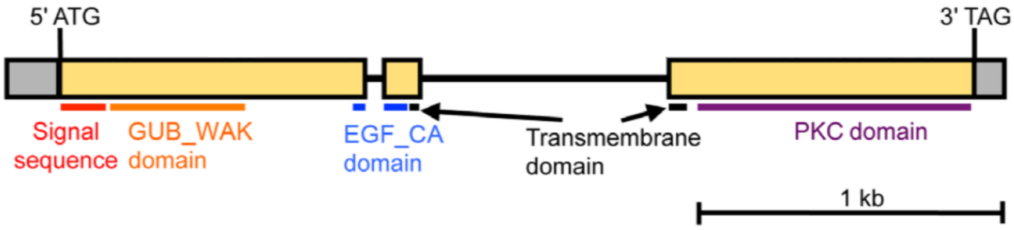


Figure 1.4 – The structure of *Snn1*. Showing exons in yellow and untranslated regions in grey. Exonic regions encoding protein domains are highlighted: GUB_WAK = galacturonic acid binding wall associated kinase, EGF_CA = calcium binding epidermal factor, PKC = protein kinase C. Adapted from Shi et al., 2016b.

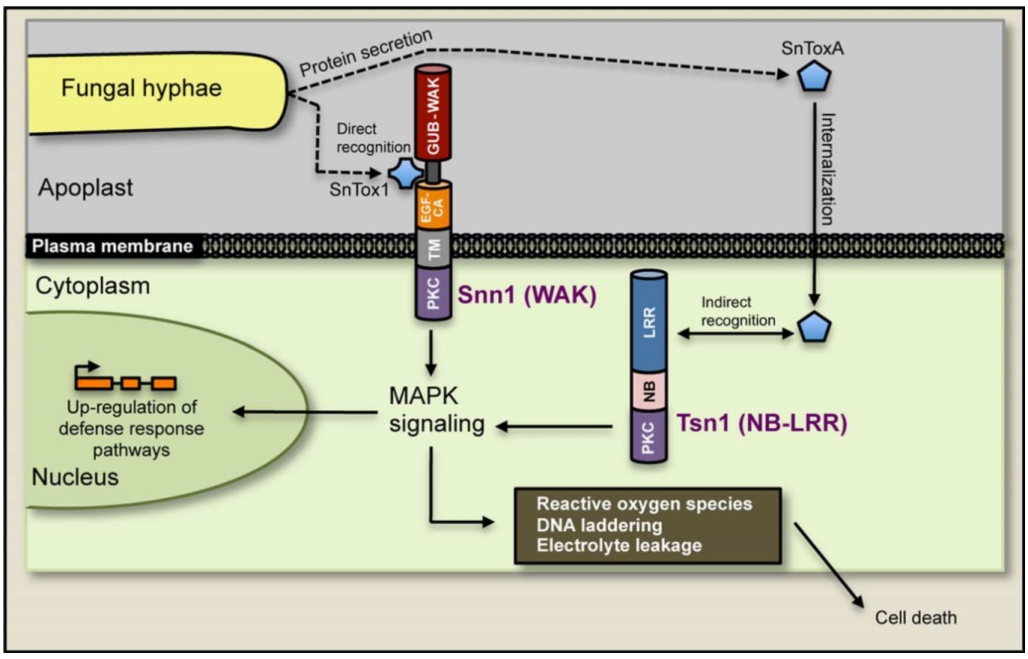


Figure 1.5 – Interaction between the fungal effectors SnTox1 and SnToxA and host proteins mediating effector sensitivity. Showing the direct interaction between SnTox1 and Snn1, the indirect interaction between SnToxA and Tsn1, the internalisation of SnToxA and the downstream triggering of Mitogen-activated protein kinase (MAPK) signaling and cell death. WAK = wall asssociated kinase, NB-LRR = nucleotide binding leucine-rich repeat, PKC = protein kinase C, TM = transmembrane domain, EGF-CA = calcium binding epidermal growth factor. Adapted from Shi et al., 2016b.

1.6.3 SnTox3-*Snn3* Interaction

The *P. nodorum* effector SnTox3 was first identified in 2008 (Friesen et al., 2008). The SnTox3 protein sequence was later characterised as a 25.8 kDa immature protein, with the first 20 of the 230 aa chain forming a signal peptide for secretion (Liu et al., 2009). SnTox3 has six cysteine residues that form disulfide bonds, at least one of these bonds is essential for biological function. Recent work has shown that an avirulent *P. nodorum* strain could become virulent with just the addition of the 693 bp intron free *SnTox3* gene (Liu et al., 2009; Waters et al., 2011).

The discovery of SnTox3 allowed for the subsequent identification of the corresponding wheat sensitivity locus, *Snn3* (more recently termed *Snn3-B1*), located on the short arm of chromosome 5B. *P. nodorum* culture filtrate containing SnTox3 was first produced using wildtype isolate, SN15, and sensitivity was mapped using the BR34 × Grandin wheat population (Friesen et al., 2008). This is congruent with data from the Australian doubled haploid mapping population Calingiri × Wyalkatchem, which was used to identify the SnTox3 sensitivity locus *QSnb.fcu-5BS* (Phan et al., 2016). However, this study also found a novel minor SnTox3 sensitivity QTL on the long arm of chromosome 4B, *QSnb.cur-4BL* (Phan et al., 2016). In 2010, an *Snn3* homoeologue was found on chromosome 5D, designated *Snn3-D1*, in a study on the bread wheat D genome donor, *Aegilops tauschii* (Zhang et al., 2010). However, *Snn3-D1* has not been reported in bread wheat. Interestingly, epistatic effects on gene expression are observed between the known *P. nodorum* effectors. For example, when *SnTox1* is knocked out, *SnTox3* expression increases in the *P. nodorum* strain SN15, thus, SnTox1 appears to be suppressing *SnTox3* expression. This is an interesting contrast to the observation that the wheat effector sensitivity loci *Tsn1* and *Snn2* are epistatic to *Snn3-B1* (Phan et al., 2016). Unlike SnToxA, it was found that SnTox3 interacts with a broader range of PR-1 proteins and it has been hypothesised that TaPR-1 proteins facilitate the infection of the host (Breen et al., 2016; Breen et al., 2017).

1.7 Aims

The SnTox3 sensitivity locus *Snn3-B1* has been recently shown to confer resistance to *P. nodorum* infection at the adult stage in the field (Ruud et al., 2017). However, while SnTox3 is a key effector controlling *P. nodorum* virulence, the SnTox3-*Snn3-B1* interaction is not understood to the same level as SnToxA-*Tsn1* or SnTox1-*Snn1*. This thesis focuses on investigating the genetics and genomics of the wheat interaction with SnTox3. The overall aims of this thesis were to:

1. Fine-map the wheat SnTox3 sensitivity locus *Snn3-B1* and identify additional SnTox3 sensitivity QTL.
2. Use RNA-sequencing to identify candidate genes for *Snn3-B1*.
3. Use molecular biological and bioinformatic techniques to further prioritise candidate genes for *Snn3-B1*.
4. Identify wheat adult plant QTL for *P. nodorum* field resistance and compare their locations with the major and minor sensitivity loci of known *P. nodorum* effectors.

2

Genetic Mapping of SnTox3 Sensitivity Loci in Wheat

2.1 Abstract

Parastagonospora nodorum is a necrotrophic fungal pathogen of bread wheat (*T. aestivum* L.), one of the world's most important crops. In *P. nodorum*, three proteinaceous effectors have been cloned: *SnToxA*, *SnTox1*, and *SnTox3*. In this chapter, sensitivity to all three *P. nodorum* effectors is surveyed in an association mapping (AM) panel of 480 European wheat varieties, identifying *Tsn1* ($-\log_{10}P \geq 31.34$) and *Snn1* ($-\log_{10}P \geq 6.41$) loci. Subsequently, the wheat SnTox3 sensitivity locus *Snn3-B1* was fine-mapped using a combination of a genome-wide association study (GWAS) in the AM panel, quantitative trait locus (QTL) analysis with a bi-parental mapping population and single marker analysis (SMA) using an eight-founder wheat multi-parent advanced generation inter-cross (MAGIC) population. GWAS identified 10 significant (Bonferroni corrected $P \leq 0.05$) markers defining a single locus, *Snn3-B1*, located on the short arm of chromosome 5B, explaining 32 % of the phenotypic variation (peak markers: Excalibur_c47452_183 and GENE-3324_338, $-\log_{10}P = 20.44$). SMA of SnTox3 sensitivity in the MAGIC population located *Snn3-B1* via five significant single nucleotide polymorphisms (SNPs), defining a 6.2 kb region that included the two peak SNPs identified in the AM panel. Accordingly, the SNP Excalibur_c47452_183 was converted to the Kompetitive Allele Specific

PCR (KASP) genotyping system, and validated by screening a subset of 95 wheat varieties, providing a valuable resource for marker assisted breeding and further genetic investigation. Additionally, QTL analysis with the bi-parental mapping population, also identified *Snn3-B1* (logarithm of the odds (LOD) = 50.82; peak marker: BS00079166_51) within a 13 cM genetic interval, explaining 57 % of the phenotypic variation. Therefore, defining a genomic region of 5.378 Mbp, with 326 kb and 6.2 kb intervals delineated by GWAS of the AM panel and SMA of the MAGIC population, respectively. The *de novo* annotation of the gene content in this region identified two candidate genes in a SnTox3 sensitive variety, which are close paralogues of *Tsn1*.

As well as the major *Snn3-B1* locus, additional minor SnTox3 sensitivity loci were identified. Composite interval mapping in the MAGIC population identified six minor SnTox3 sensitivity QTL, on chromosomes 2A (*QTox3.niab-2A.1*, -$\log_{10}$P = 6.74), 2B (*QTox3.niab-2B.1*, -$\log_{10}$P = 1.74), 3B (*QTox3.niab-3B.1*, -$\log_{10}$P = 6.74), 4D (*QTox3.niab-4D.1*, -$\log_{10}$P = 1.55), 6A (*QTox3.niab-6A.1*, -$\log_{10}$P = 3.72), and 7B (*QTox3.niab-7B.1*, -$\log_{10}$P = 1.70), each accounting for 3.1–6.0 % of the phenotypic variation. In the Avalon × Cadenza population, minor QTL were identified on chromosomes 1B (*QTox3.niab-1B*, LOD = 6.11) and 2B (*QTox3.niab-2B.2*, LOD = 3.83).

Collectively, the phenotypic, genetic and genomic datasets presented here provide wheat researchers and breeders with knowledge and resources regarding the sensitivity of European wheat germplasm to *P. nodorum* effectors, as well as simple diagnostic markers for predicting allelic state at *Snn3-B1*.

2.2 Introduction

The necrotrophic pathogen *Parastagonospora nodorum*, the causal agent of the disease SNB and glume blotch in wheat, is of significant economic importance in Australia, Europe, North America, and Northern Africa (Friesen et al., 2005; Oliver et al., 2012; Quaedvlieg et al., 2013). The visual symptoms of chlorosis and necrosis of the leaf tissue and glumes, often in the form of lesions, resulting in grain yield losses of up to 31 % (Bhathal et al., 2003; Solomon et al., 2006). *P. nodorum* is thought to derive nutrients from dead plant tissue, utilising fungal effector proteins, to induce a hypersensitive response in the host, which takes the form of programmed cell death (Friesen et al., 2007; Liu et al., 2009; Oliver et al., 2012). The necrotic response in the sensitive host is hypothesised to facilitate pathogen colonisation, promoting infection and ultimately providing a rich nutrient source (Oliver and Solomon, 2010; Vincent et al., 2012). This is known as effector-triggered susceptibility (ETS) and is genetically induced via an 'inverse gene for gene system' (Friesen et al., 2007).

The identification of effector proteins has led to a paradigm shift in the approach to tackling these types of pathogens, as the host–pathogen interactions can be broken down into their constituent parts. Consequently, targeted breeding could then be used to eliminate host sensitivity on an effector by effector basis.

Effector proteins were described for the first time with regards to a host–pathogen interaction between the necrotroph *Alternaria alternata* and *Pirus serotine* (Tanaka, 1933). However, the first protein effector described in a necrotrophic pathogen was PtrToxA from the wheat tan spot pathogen, *Pyrenophora tritici-repentis*, which triggers necrosis in wheat lines carrying susceptible alleles at the *Tsn1* locus (Ballance et al., 1989; Tomas et al., 1990; Faris et al., 1996). A near identical effector, SnToxA, was discovered in *P. nodorum*, with the corresponding host sensitivity locus also being Tsn1 (Liu et al., 2006). *Tsn1*, located on the long arm of chromosome 5B, has been cloned and encodes an intracellular protein with a serine/threonine protein kinase (S/TPK) domain, a nucleotide-binding site (NBS), and leucine-rich repeats (LRRs), with deletion of *Tsn1* resulting in SnToxA insensitivity (Faris et al., 2010). Similarly, the purification and subsequent isolation of the *P. nodorum* effector, SnTox1, allowed identification of the corresponding wheat sensitivity locus, *Snn1*, on the short arm of chromosome 1B (Liu et al., 2004; Liu et al., 2012; Cockram et al., 2015). Map-based cloning found *Snn1* to encode a wall-associated kinase (WAK), with yeast two-hybrid analysis showing that Snn1 and SnTox1 proteins interact directly *in vitro*, unlike Tsn1 and SnToxA (Faris et al., 2010; Shi et al., 2016b). Given the nature of their corresponding wheat sensitivity loci, it is hypothesised that SnToxA and SnTox1 activate the wheat pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) and effector-triggered immunity (ETI) pathways respectively, which for biotrophic pathogens protect against pathogen infection. However, as *P. nodorum* is a necrotrophic pathogen, the triggering of these pathways, which induce necrosis and cell death, promotes *P. nodorum* growth and propagation (Shi et al., 2016b). Wheat varieties carrying both *Tsn1* and *Snn1* show higher levels of necrosis upon *P. nodorum* infection than those varieties carrying either *Tsn1* or *Snn1* alone, indicating that the hijacking of both the PTI and ETI pathways for necrotrophic effector triggered susceptibility supports pathogen survival and reproduction (Chu et al., 2010; Shi et al., 2016b).

Characterisation of a third *P. nodorum* effector, SnTox3, led to the identification of its corresponding wheat sensitivity locus, *Snn3* (more recently termed *Snn3-B1*), located on the short arm of chromosome 5B. Culture filtrate containing SnTox3 was produced using a wildtype pathogen isolate, SN15, and host sensitivity was mapped using the BR34 × Grandin wheat population, accounting for 17 % of the phenotypic variation (Friesen et al., 2008). This agrees with data from the doubled haploid mapping population, Calingiri × Wyalkatchem, which identified the *SnTox3*

sensitivity locus *QSnb.fcu-5BS*, as well as a minor SnTox3 sensitivity QTL on the long arm of chromosome 4B, *Qsnb.cur-4BL* (Phan et al., 2016).

2.2.1 Aims

Understanding the effector sensitivities of wheat varieties, and the genetic determinants controlling wheat sensitivity to *P. nodorum*, allows informed manipulation of alleles and germplasm within wheat breeding programs. In this chapter, an AM panel of 480 European wheat varieties is surveyed for sensitivity to SnToxA, SnTox1, and SnTox3 and subsequently used in concert with a MAGIC population and the Avalon × Cadenza (A×C) doubled haploid population to fine-map *Snn3-B1*, and to identify additional minor QTL for SnTox3 sensitivity, having used the SnToxA and SnTox1 data and the known locations for *Tsn1* and *Snn1* to validate the methodology.

2.3 Materials and Methods

2.3.1 Wheat Germplasm and High-Density Genotyping

Three bread wheat (*T. aestivum* L.) genetic mapping populations were used for effector sensitivity screening and genetic mapping. The first was an AM panel, representing a diverse collection of 480 wheat varieties drawn from historic collections and National Lists, encompassing varieties released between 1916 and 2007. Of these, 420 were released or marketed within the UK; however, many of these were bred for initial release outside of the UK. The remaining 60 varieties do not have a UK Application For Protection (AFP) number, and therefore where country of origin information was not available, these varieties were assumed to either be non-UK, or to represent accessions that predate the application process. The majority of the AM panel represents UK varieties (330 lines, 68.8 % of the total collection), followed by 51 French (10.6 %), 37 German (7.7 %), and 19 Dutch varieties (4.0 %). The remaining 17 varieties (3.5 %) with country information come from Belgium, Canada, Denmark, Sweden, Switzerland, and the United States of America. For 26 varieties (5.4 %) country of origin was not known. The population was previously genotyped using the iSelect 90,000 feature wheat SNP array (Illumina, San Diego, CA; Wang et al., 2014), resulting in 26,018 polymorphic SNPs with a minor allele frequency ≤ 3 % (data available via https://www.niab.com/pages/id/491/Wheat_90k_SNP_dataset). The second was the bi-parental A×C doubled haploid population, used for effector sensitivity screening and genetic mapping. The population consists of 206 lines (Ma et al., 2015), of which 198 lines were used in this chapter. The population was previously genotyped using the iSelect 90,000 feature wheat SNP array, resulting in

10,941 polymorphic SNPs. 9,266 SNPs are present in the A×C map (see Chapter 4 for map construction). The third was an eight-founder 'NIAB Elite MAGIC' population (Mackay et al., 2014), the founders of which (cvs. Alchemy, Brompton, Claire, Hereward, Rialto, Robigus, Soissons and Xi19) were selected for their high seed yield, disease resistance, and their range of end-use qualities. The founders were intercrossed in a simple replicated funnel crossing scheme over three generations, with individuals from the eight-way families subsequently selfed over four generations through single seed descent to produce > 1,000 recombinant inbred liness (RILs). A subset of these F_5 RILs were genotyped using the iSelect 90,000 feature SNP array detailed above, resulting in 20,643 polymorphic markers (Gardner et al., 2016). These data allowed the development of a high-resolution genetic map consisting of 18,601 markers mapped using 643 MAGIC lines (Gardner et al., 2016). The remaining 2,042 SNPs were not mappable, due largely to segregation distortion and/or dominance (Gardner et al., 2016).

2.3.2 Effector Protein Production and Wheat Phenotyping

P. nodorum effectors were supplied via co-supervisor Richard Oliver, Kar-Chun Tan and his group at Curtin University, Australia. In brief, SnTox1 and SnTox3 were expressed in *Pichia pastoris*, as previously described (Tan et al., 2014). For SnToxA, heterologous expression was conducted in *Escherichia coli* BL21E using the pET21a expression vector as described in Tan et al. (2012). Protein preparations were desalted in 20 mM pH 7.0 sodium phosphate, freeze-dried for storage, and subsequently re-suspended in ultra-pure water and stored at 4 °C prior to use. The AM, A×C and MAGIC lines were grown in 96-well trays with fine/medium compost (M3) in a heated and lit glasshouse 20 °C/17 °C day/night with a 16 hour photoperiod. Each line was represented by three to four replicates, and each MAGIC founder by eight replicates, with experimental design carried out using MATLAB (MATLAB, The MathWorks Inc., Natick, Massachusetts, United States) or R/blocksdesign. For the AM panel, randomisation was performed using a custom software routine written in the MATLAB programming environment, to include three biological replicates of each line. For the A×C and MAGIC populations, the experimental design was split into four blocks, each block containing one replicate of each line and two replicates of each of the parents, with line positions randomised within each block, using R/blocksdesign. Therefore, a total of four biological replicates of each line and eight biological replicates of each parent were included. Infiltration on the AM, A×C and MAGIC populations was undertaken on seedlings at Zadoks growth stage (GS) 12 (Zadoks et al., 1974), as previously described (Tan et al., 2012). Briefly, a 1 ml plastic syringe was used to infiltrate approximately 50 µl of either SnToxA, SnTox1,

or SnTox3 suspension into the first leaf, with the extent of the leaf infiltration region marked with a non-toxic pen. Seven days following infiltration, the plants were visually evaluated for effector sensitivity on a scale of 0 (insensitivity, no symptoms) to 4 (extensive necrosis; Figure 2.1; Tan et al., 2012). A water control was also used to establish a symptom baseline to evaluate possible damage due to the infiltration process. Mean sensitivity scores were calculated per line for subsequent analysis. The AM panel phenotyping and trial design was carried out by a team at NIAB prior to the start of the PhD project.

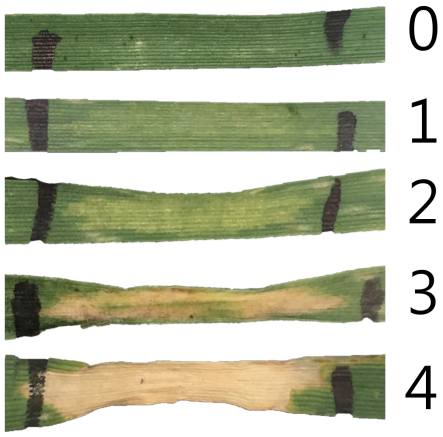


Figure 2.1 – Visual symptoms of host sensitivity to *P. nodorum* effectors and their corresponding score. 0 = insensitive, 1 = slight chlorosis, 2 = chlorosis, 3 = chlorosis with some necrosis, 4 = necrosis.

2.3.3 Statistical Analyses and Bioinformatics

Effector phenotypic data were analysed using R (R Core Team, 2013) to determine mean sensitivity scores for each variety as well as variance within each variety. For each effector, the heritability of line means (broad sense) was calculated by first estimating components of variation from ANOVA (in the AM panel) or REML (in MAGIC) while taking into account all features of the experimental designs. Heritability was then estimated as

$$h^2 = \sigma^2{}_{\mathrm{G}}/(\sigma^2 + \sigma^2 e)$$

where $\sigma^2{}_G$ is the genetic variation between line means and σ^2e is the error variance appropriate to those means. Calculations for the AM panel and MAGIC population were carried out in GenStat (v19; VSN International, 2011) and the package lme4 (Bates et al., 2015) in R. Heritability calculations for the A×C population were carried out in IciMapping (v3.3; Meng et al., 2015). A GWAS using the AM panel

was undertaken using the efficient mixed model association (EMMA) algorithm (Kang et al., 2008) using a compressed mixed linear model (CMLM; Zhang et al., 2010) which includes both fixed and random effects, implemented with the Genome Association and Prediction Integrated Tool (GAPIT) package (Lipka et al., 2012) in R. For GWAS, Bonferroni corrected P = 0.05 and P = 0.01 significance thresholds were used, termed here 'significant' and 'highly significant', respectively. For MAGIC, correction for multiple testing was carried out using R/qvalue using a corrected P-value threshold of 0.05. MAGIC genetic analyses were undertaken using two approaches.

1. SMA: a simple linear model test in R/lme4 using all 20,643 SNPs. After finding of a major QTL, the analysis was repeated with the major QTL as a covariate.

2. Haplotype analysis using a subset of 7,369 uniquely mapped SNPs from the MAGIC genetic map (Gardner et al., 2016).

Founder haplotype probabilities were computed with the "mpprob" function in R/mpMap (Huang and George, 2011), implemented in R/qtl (Broman et al., 2003), using a threshold of 0.05. QTL analysis with haplotypes was carried out using (a) linear mixed model using all mapped markers, (b) interval mapping (IM) using the mpIM function in R/mpMap, and (c) composite interval mapping (CIM) with either 5 or 10 covariates in R/mpMap. The function sim.sigthr in R/mpMap was used to conduct 100 simulations using the dataset to obtain an empirical QTL significance threshold at a P-value threshold of 0.05. A full QTL model was then fitted with all QTL using R/fit.mpQTL. For the AM panel, the difference between homozygous marker classes was estimated as twice the linear regression coefficient from a regression of the trait on the marker classes (coded 0, 1, and 2 with 1 being the heterozygous genotype). The coefficient of determination (or r-squared) was used as a measure of the proportion of variance explained by the marker. Genetic markers were anchored using BLASTn (Altschul et al., 1990) against the wheat cv. Chinese Spring 42 IWGSC RefSeq assembly (v1.0; Appels et al., 2018), and where explicitly stated in the text, against the cv. Chinese Spring 42 TGAC assembly (v1.0; Clavijo et al., 2017). In the case of hits with an equal match on multiple homoeologues, chromosome allocation followed that assigned by the MAGIC genetic map (Gardner et al., 2016), where possible. Nomenclature for the QTL discovered in this study follows that recommended by the Catalog of Gene Symbols for Wheat (McIntosh et al., 2008). Predicted protein domains were identified using Pfam 31.0 (Finn et al., 2016). QTL analysis for the A×C population was carried out using the additive mapping algorithm in IciMapping and the linkage map developed in Chapter 4. A LOD threshold was calculated using 1000 permutations with a 0.05

level threshold. The QTL intervals are defined using the two outer most flanking markers that are still significant. Using BLASTn (Altschul et al., 1990) cv. Cadenza assembly scaffolds (EIv1; earlham.ac.uk/grassroots-genomics) were anchored to a physical location on the IWGSC RefSeq assembly. This enabled identification of the Cadenza scaffolds as being located within the *Snn3-B1* QTL region. FgeneSH (Salamov and Solovyev, 2000) was used to *de novo* annotate the identified Cadenza scaffolds and then the Conserved Domain Database (Marchler-Bauer et al., 2010) was used to identify protein domains for the gene annotation.

2.3.4 KASP Marker Development

Selected SNPs were converted to the KASP genotyping platform (LGC Genomics, Hoddesdon, UK). Genomic DNA sequence flanking the target SNPs were used to design KASP primers using the software PolyMarker (Ramirez-Gonzalez et al., 2015). Genomic DNA was extracted from seedling leaves harvested from a subset of the AM panel using a modified Tanksley protocol (Fulton et al., 1995), DNA concentrations were determined using a Nanodrop 200 spectrophotometer (Thermo Scientific, Waltham, MA), and diluted to a final concentration of $10\,\mathrm{ng}\cdot\mu\mathrm{l}^{-1}$ using sterile PCR-grade water. KASP design and genotyping was undertaken by a service provider following the manufacturer's guidelines (LGC Genomics, Hoddesdon, UK), returned as csv files, analysed with SNP Viewer (v1.99; LGC Genomics, Hoddesdon, UK), and compared against the corresponding SNP calls from the iSelect 90,000 feature SNP array.

2.3.5 Candidate Gene Presence/Absence PCR Assay

A multiplex presence / absence polymerase chain reaction (PCR) assay was designed for the two candidate genes identified in this chapter (termed 'Gene1' and 'Gene2'). Genomic DNA sequences for Gene1 and Gene2 from cv. Cadenza were aligned using Clustal Omega (Madeira et al., 2019), followed by manual design of a common forward primer: 5′-CACTGAAGGGGAAGATGTCTGA-3′. Reverse primers were then manually designed specific to Gene1 (5′-AACATGGGCGTGTGCACTAT-3′) and Gene2 (5′-GGTGAATCCAACGGCATCCA-3′), predicted to PCR products from genomic DNA template of 456 base pairs (bp) and 238 bp respectively. For each line of the AM panel, PCR amplification was undertaken in a total volume of 20 µl containing 5 ng of genomic DNA mixed with 0.2 µl FastStart *Taq* DNA polymerase ($5\mathrm{u}\cdot\mu\mathrm{l}^{-1}$; Roche), 2.5 µl FastStart ×10 buffer without $MgCl_2$, 2.5 µl ($25\,\mathrm{mmol}\cdot\mathrm{dm}^{-3}$) $MgCl_2$, 0.5 µl dNTP ($10\,\mathrm{mmol}\cdot\mathrm{dm}^{-3}$ each), 1 µl ($10\,\mu\mathrm{mol}\cdot\mathrm{dm}^{-3}$) of each of the reverse primers and 2 µl ($10\,\mu\mathrm{mol}\cdot\mathrm{dm}^{-3}$) of the common primer. PCR reactions were incubated at 96 °C for 60 s followed by 40 cycles of 50 s at 96 °C, 50 s

at 67 °C and 60 s at 72 °C, followed by a final incubation 72 °C for 240 s. All PCRs were carried out using a Veriti 96 well thermal cycler (Applied Biosystems, Waltham, MA). Primers known to amplify *GAPDH* (5′-TTACGACTTGCGAAGCCAGCA-3′ and 5′-AAATGCCCTTGAGGTTTCCC-3′) were used in separate PCRs to serve as a positive control, thus helping to identify possible false negative PCR result for Gene1 and Gene2. All PCRs were undertaken using at least two technical replicates per line and reaction products were visualised by agarose gel electrophoresis.

2.4 Results

2.4.1 *P. nodorum* Effector Sensitivity Phenotyping

The AM panel, consisting of 480 varieties and breeding lines released between 1916 and 2007, was phenotyped for sensitivity to SnToxA, SnTox1, and SnTox3 via leaf infiltration at the seedling stage, and the severity of host response scored using a 0 (insensitive) to 4 (extensive necrosis) scale (Figure 2.2). Broad sense heritability for effector sensitivity was found to be highest for SnTox3 ($h^2 = 0.92$), followed by SnToxA ($h^2 = 0.83$) and SnTox1 ($h^2 = 0.77$). For SnToxA, separation into insensitive/weakly sensitive ($0 \leq$ score 1) and strongly sensitive (score ≥ 3) was observed, accounting for 73 % (338/460 varieties) and 10 % (47/460 varieties) of all lines successfully screened, respectively (Figure 2.2). While spring cultivars represent just 6 % of the complete AM panel, 23 % (11/47 cultivars) of all SnToxA sensitive lines are spring cultivars, as is the oldest SnToxA sensitive variety, Garnet (released in 1926). Response to SnTox1 is more evenly spread across sensitivity classes, with 24 % insensitive (111/461 varieties), 28 % sensitive (130/461), and 48 % (220/461) showing intermediate sensitivity ($1 \leq$ score < 3; Figure 2.2). The most common effector sensitivity identified in the AM panel was to SnTox3, with 42 % of varieties found to be strongly sensitive (192/457), including the recent varieties Cocoon, KWS Santiago, Orator, Rainbow, and Tuxedo, all released in 2007. Insensitivity to SnTox3 was found in 25 % (114/457) of lines, including six released in 2007, while intermediate sensitivity was observed in 33 % of varieties. All possible combinations of sensitivity between the three effectors were identified, with 18 varieties displaying high sensitivity (score ≥ 3) to all three effectors.

2.4.2 Genetic Analysis of *P. nodorum* Effector Sensitivity using the AM Panel

Initially, the precision of the AM panel was assessed empirically by undertaking GWAS for SnToxA sensitivity, known to be controlled by allelic variation at the *Tsn1*

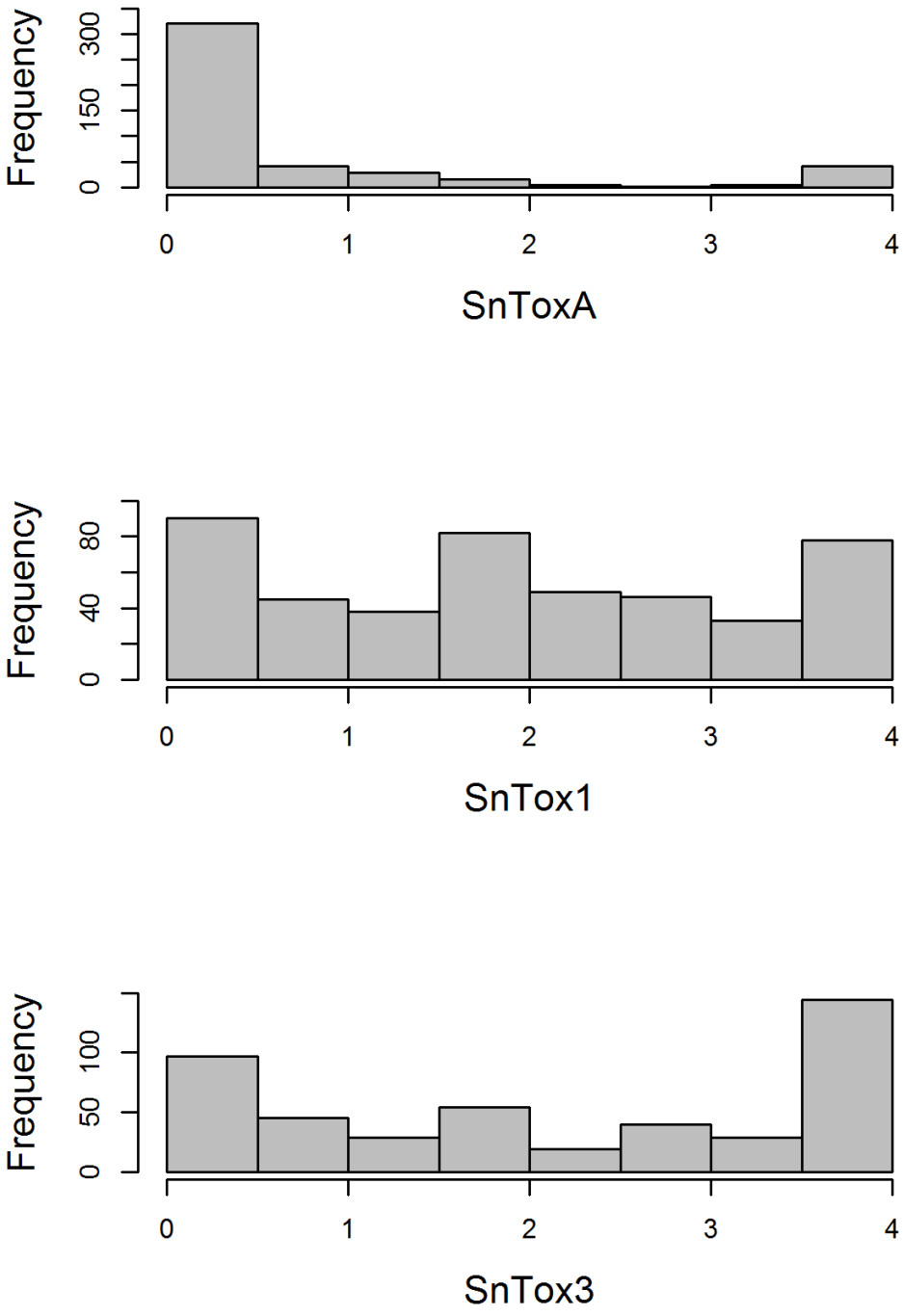


Figure 2.2 – Histogram of host sensitivity to *P. nodorum* effectors in a wheat association mapping panel of 480 northwest European varieties. Showing the frequency of sensitivity scores (0 = insensitive, 4 = sensitive) to SnToxA, SnTox1 and SnTox3 for each accession, represented by a mean of four biological replicates.

locus. Using a data matrix of 26,018 SNPs across 480 varieties, and a Bonferroni corrected P = 0.01 significance threshold (-$\log_{10}P$ = 6.41), GWAS identified 30 highly significant markers associated with SnToxA sensitivity. These accounted for between 25 % and 60 % of the total variation (average 38 %) with differences in score between the two homozygous classes ranging from 1.5 to 3.1 (mean = 2.5). Of the eight most significant markers (-$\log_{10}P \geq 31.34$), six are located within a gene model encoding a potassium transporter (TraesCS5B01G368500), just two genes proximal to the serine/threonine protein kinase (S/TPK)-nucleotide binding site (NBS)-leucine-rich repeat (LRR) gene underlying *Tsn1* (TraesCS5B01G059000; Faris et al., 2010). Similarly, GWAS of SnTox1 sensitivity identified seven highly significant (-$\log_{10}P \geq 6.41$) SNPs. These accounted for between 8.3 % and 14.3 % of the variation (average 11.6 %) with differences in score between the two homozygous classes ranging from 0.9 to 1.6 (mean = 1.4). All seven SNPs were located at the Snn1 locus on chromosome 1B, with SNP Excalibur_c21898_1423 located 25 genes distant from the WAK gene underlying *Snn1* in cv. Chinese Spring (TraesCS1B01G004100).

Having demonstrated the utility of the AM panel, the SnTox3 sensitivity phenotypic data was used to undertake GWAS, identifying 14 significant SNPs (Table 2.1 and Figure 2.3). Of these, seven were located within a single region on the short arm of chromosome 5B in the IWGSC RefSeq assembly. The most significant SNPs were Excalibur_c47452_183 and GENE-3324_338 (-$\log_{10}P$ = 20.44). These explained 32 % of the phenotypic variation, with a phenotypic difference between homozygous allele classes of 2.1. Anchoring previous markers (Ruud et al., 2017) identified as flanking *Snn3-B1* to the IWGSC RefSeq assembly (v1.0) confirmed that *Snn3-B1* had been identified in the AM panel. Using the most significant *Snn3-B1* markers identified in the AM panel (Excalibur_c47452_183) as a co-factor in GWAS found no additional significant genetic loci.

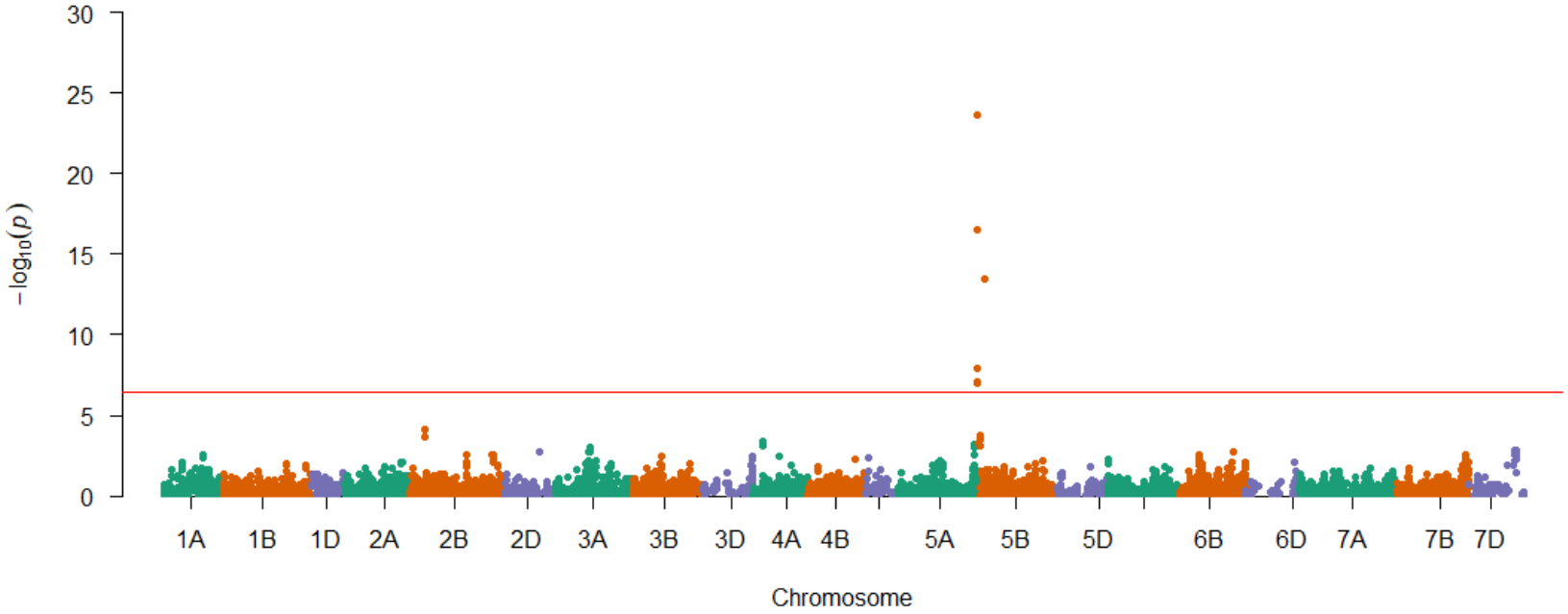


Figure 2.3 – Genetic mapping of SnTox3 sensitivity in the AM panel. SNPs are ordered according to the genetic map (Gardner et al., 2016). Significant SNPs that were unmapped in the genetic map are included on the relevant chromosome (5B) via their physical map positions (IWGSC RefSeq v1.0), relative to those of the genetically mapped chromosome 5B SNPs. The Bonferroni corrected P = 0.01 significant threshold is indicated in red ($-\log_{10}$P = 6.41).

Table 2.1 – Significant SNPs (P ≤ 0.05, Bonferroni corrected) for SnTox3 sensitivity identified in **(A)** the AM panel and **(B)** single marker analysis on 20,643 SNPs in the MAGIC population, listing the SNPs with predicted allelic effect > 1.0.

SNP	MAF	-$\log_{10}P$	Chr, SNP position (bp)	IWGSC RefSeq v1.0 gene model	Gene annotation
A					
Excalibur_c47452_183	0.194	23.62	5B, 6654166	TraesCS5B01G005100	Ubiquitin conjugating enzyme E2
GENE-3324_338	0.194	23.62	5B, 6647920†	TraesCS5B01G005000	Nucleotide triphosphate hydrolase
BobWhite_c4838_58	0.194	23.62	5B, 6654053	TraesCS5B01G005100	Ubiquitin conjugating enzyme
BS00091518_51	0.194	23.62	5B, 6648547	TraesCS5B01G005000	Nucleotide triphosphate hydrolase
BS00091519_51	0.194	23.62	5B, 6648567	TraesCS5B01G005000	Nucleotide triphosphate hydrolase
BS00064297_51b	0.272	16.52	5B, 6974807	TraesCS5B01G005600	Transmembrane protein, putative (DUF594)
BS00064298_51b	0.272	16.52	5B, 6974825	TraesCS5B01G005600	Transmembrane protein, putative (DUF594)
Ex_c1846_1818b	0.449	14.87	5B, 64632597	TraesCS5B01G059000	Transmembrane protein, putative (DUF594)
Ex_c1846_1818a	0.271	13.46	5B, 64736555	TraesCS5B01G059000	Transmembrane protein, putative (DUF594)
RAC875_c7582_680	0.124	7.88	5B, 2058821	TraesCS5B01G002000LC	None
Ku_c10387_272	0.124	7.88	5B, 232228	TraesCS5B01G0006000	Microtubule-associated protein 70-2
RAC875_c39204_91	0.157	7.11	5B, 6852650	None	None
BS00064297_51a	0.156	7.02	5B, 6974807	TraesCS5B01G0005600	Transmembrane protein, putative (DUF594)
BS00064298_51a	0.156	7.02	5B, 6974825	TraesCS5B01G0005600	Transmembrane protein, putative (DUF594)
B					
Excalibur_c47452_183		15.7‡	5B, 6654166	TraesCS5B01G0005100	Ubiquitin conjugating enzyme E2
GENE-3324_338		15.7‡	5B, 6647920†	TraesCS5B01G0005000	Nucleotide triphosphate hydrolase
Bobwhite_c4838_58		15.7‡	5B, 6654053	TraesCS5B01G0005100	Ubiquitin conjugating enzyme E2
BS00091518_51		15.7‡	5B, 6648547	TraesCS5B01G0005000	Nucleotide trisphosphate hydrolase
BS00091519_51		15.7‡	5B, 6648547	TraesCS5B01G0005000	Nucleotide trisphosphate hydrolase

MAF, minor allele frequency (AM panel only); Chr, chromosome; U, chromosome unknown.

† While chromosome 5A homoeologues represented a better BLASTn hit, details for the 5B homoeologues presented here, as this SNP has previously been shown to be located on chromosome 5B using the SNP as a trait, and localising to a chromosome by trait mapping (Gardner et al., 2016).

‡ P-values < 2.2^{-16}, so -$\log_{10}$P set arbitrarily to > 15.7 here.

2.4.3 Genetic Analysis of SnTox3 Sensitivity using the Avalon × Cadenza Population

The wheat A×C doubled haploid population was phenotyped for sensitivity to SnTox3 via leaf infiltration (n = 198), and the severity of host response scored using a 0 (insensitive) to 4 (extensive necrosis) scale. Heritability for SnTox3 sensitivity was found to be high ($h^2 = 0.71$).

In the AM panel, Avalon and Cadenza were found to contrast for SnTox3 sensitivity with scores of 0.8 and 4, respectively. SnTox3 sensitivity scores in the A×C progeny ranged from 0 to 4, with the majority of lines displaying intermediate sensitivity (1 ≥ score ≥ 3, 56.6 % lines) or no sensitivity (score < 1, 36.4 % lines) with a mean score of 1.55. QTL analysis was carried out using the improved A×C genetic map developed in Chapter 4, identifying three QTL on chromosomes 1B, 2B and 5B (Figure 2.2). The first QTL, on chromosome 1B, termed here *QTox3.niab-1B*, was located at 68.48 cM (349.715 Mbp) by SNP Tdurum_contig76507_422 (LOD = 6.11, percentage variation explained (PVE) = 4.41) between Ra_c11303_583 (67.72 cM) and Tdurum_contig13335_194 (73.8 cM). The second, at 240.39 cM on chromosome 2B (*QTox3.niab-2B.2*, LOD = 3.83, PVE = 2.55), was identified by SNP Excalibur_c30571_95 (771.205 Mbp), between RAC875_rep_c108836_55 (235.05 cM) and BS00062699_51 (247.26 cM). The third (*Snn3-B1*) was identified by SNP BS00079166_51 (7.706 Mbp) on chromosome 5B at 45.17 cM showing the highest significance (LOD = 50.82, PVE = 57.27), between BS00024993_51 (44.66 cM) and BS00032003_51 (58.15 cM).

Table 2.2 – Avalon × Cadenza QTL for SnTox3 sensitivity

QTL	Chr, cM (Mbp)	Most significant SNP	LOD	PVE
QTox3.niab-1B	1B, 68.48 (349.715)	Tdurum_contig76507_422	6.11	4.41
QTox3.niab-2B.2	2B, 240.39 (771.205)	Excalibur_c30571_95	3.83	2.55
Snn3-B1	5B, 45.17 (7.706)	BS00079166_51	50.82	57.27

SnTox3 sensitivity QTL discovered in the Avalon × Cadenza doubled haploid population. LOD = logarithm of the odds. PVE = percentage variation explained.

2.4.4 Genetic Analysis of SnTox3 Sensitivity using MAGIC

While screening the AM panel for SnTox3 sensitivity, the eight founders of the NIAB Elite MAGIC population were found to contrast for SnTox3 sensitivity. Accordingly, 643 lines of the MAGIC population (3–4 reps/line), as well as the eight founders (8 reps/line), were subsequently phenotyped for SnTox3 sensitivity. The sensitivity scores of the parents ranged from 0 (cvs. Alchemy, Claire, Robigus) to 4 (cvs. Hereward, Rialto, Soissons, Xi19), with cv. Brompton displaying intermediate sensitivity (1.93). SnTox3 sensitivity scores in the MAGIC progeny ranged from 0 to

4, with the majority of lines displaying high sensitivity (score ≥ 3.5, 42.5 % lines) or no sensitivity (score ≤ 0.5, 27.5 % lines). Heritability for SnTox3 sensitivity in the MAGIC population was calculated to be very high ($h^2 = 0.95$). Initially, the 643 MAGIC lines along with the 20,643 mapped and unmapped SNPs were used for SMA using a simple linear model test, identifying 114 significant ($P < 0.05$) markers. Of these, the five most significant MAGIC SNPs, with predicted allelic effects > 1, are the same as the five most significant markers identified in the AM panel (Excalibur_c47452_183, GENE-3324_338, BobWhite_c4838_58, BS00091518_51, and BS00091519_51; Table 2.1). SMA identified four additional QTL. The first was on chromosome 2B (termed here *QTox3.niab-2B.1*, P = 0.023), located at 356.66 cM by SNP Kukri_c9898_1766. The second, at 40.61 cM on chromosome 4D (*QTox3.niab-4D.1*, P = 0.037), was identified by SNP BS00036421_51, the third (*QTox3.niab-6B.1*) by three SNPs on chromosome 6B, with wsnp_Ku_c2119_4098330 showing the highest significance (P = 0.003), and the fourth (*QTox3.niab-7B.1*) by four chromosome 7B SNPs, with BS00022127_51 showing the highest significance (P = 0.038). Using the most significant SNP GENE_3324_338 at the *Snn3-B1* locus as a covariate in SMA did not identify any additional genetic loci, and resulted in the disappearance of all four minor QTL. Additionally, IM and CIM using 5 and 10 covariates was undertaken, using the 7,369 uniquely mapped SNPs from the MAGIC genetic map (Gardner et al., 2016). As well as identifying *Snn3-B1* on chromosome 5B (P-values for all analyses ≤ 2.2^{-16}, accounting for ≥ 16.95 % of the phenotypic variation), IM/CIM also identified two QTL found using SMA: *QTox3.niab-2B.1* (P = 0.047, 2.3 % variation explained, detected with IM) and *QTox3.niab-7B.1* (P = 0.025, 1.6 % variation explained, detected with IM). In addition, three further QTL were discovered, distinct to those identified by SMA (Table 2.3). The first, *QTox3.niab-2A.1*, mapped to chromosome 2A at 234.62 cM (SNPs: BS00070979_51 and Excalibur_c20478_641; positioned at ~758 Mb) with a P-value = 9.17^{-7}, and explained 6.0 % of the phenotypic variation using IM. The second, *QTox3.niab-3B.1*, was located on chromosome 3B at 84.11 cM (P = 48.51^{-4}; SNPs: wsnp_Ex_c11246_18191331, wsnp_Ex_c22401_31592784, ~68 Mb), and explained 3.1 % of the variation (found via CIM using 5 or 10 covariates). Finally, *QTox3.niab-6A.1* at 65.6 cM on chromosome 6A explained 4.2 % of the variation, identified using IM (P = 8.51^{-4}; SNPs: BobWhite_c13839_135 and IACX7801, ~22 Mb).

2.4.5 Development of KASP Genetic Markers for *Snn3-B1*

One of the most significant SNP identified in both the AM panel by GWAS and MAGIC population by SMA was Excalibur_c47452_183. This marker was selected for conversion from the iSelect 90,000 feature SNP array to the KASP genotyping

Table 2.3 – Significant SnTox3 sensitivity QTL identified by IM and/or CIM in the MAGIC population. For CIM, 5 and 10 covariates were used in the analysis (labeled below as cov 5, and cov 10, respectively).

IM		
Chr, cM	***P*-value, PVE**	**SNPs flanking QTL peak**
2A, 234.62	9.17^{-7}, 6.0	BS00070979_51, Excalibur_c20478_641
2B, 379.61	4.7^{-3}, 2.3	BS00064483_51, Kukri_c1526_666
5B, 1.27	3.33^{-16}, 15.5	BS00015136_51, GENE-3277_145
6A, 65.6	6.16^{-5}, 4.2	Bobwhite_c13839_135, IACX7801
7B, 9.43	0.025, 1.6	BS00022127_51, Kukri_c67849_109
CIM Cov 5		
2A, 235.12	1.18^{-5}, 4.9	Excalibur_c20478_641, Tdurum_contig56321_232
3B, 84.11	8.51^{-4}, 3.1	wsnp_Ex_c11246_18191331, wsnp_Ex_c22401_31592784
5B, 3.29	0, 17.0	BS00025784_51, BS00065732_51
CIM Cov 10		
2A, 234.62	3.22^{-6}, 2.8	BS00070979_51, Excalibur_c20478_641
3B, 84.11	7.96^{-4}, 3.1	wsnp_Ex_c11246_18191331, wsnp_Ex_c22401_31592784
5B, 3.29	0†, 17.0	BS00025784_51, BS00065732_51
6A, 65.6	3.53^{-4}, 3.2	BobWhite_c13839_135, IACX7801

PVE = percent variation explained. QTL identified by two or more IM/CIM analysis methods are reported in the text, using P-values and SNP information from the analysis with the least number of covariates. Chromosome (Chr) and centimorgan (cM) positions are from the NIAB Elite MAGIC genetic map (Gardner et al., 2016).

† $< 2.2^{-16}$.

platform, a single-plex technology that allows flexible, low-cost use for marker-assisted breeding and research. Primers were designed and tested on a subset of 95 varieties from the AM panel (Figure 2.4). Comparison of Excalibur_c47452_183 allele calls from KASP genotyping with those returned by the iSelect 90,000 array genotyping of the AM panel found perfect correspondence between the two, indicating robust conversion to the KASP genotyping platform. This SNP provides good, but not perfect, prediction of SnTox3 sensitivity phenotype in the AM panel (Table 2.4).

2.4.6 Analysis of the *Snn3-B1* Physical Region

To investigate gene content at the *Snn3-B1* locus on chromosome 5B, the flanking DNA sequences of the most significant SNPs identified in the AM panel ($-\log_{10}P > 16$,

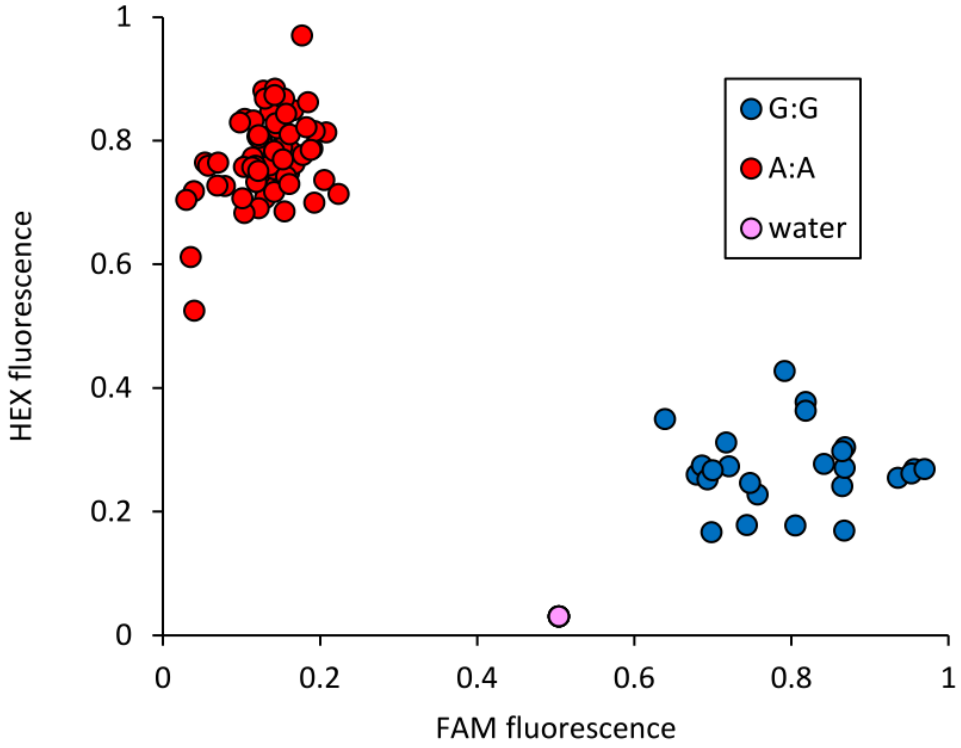


Figure 2.4 – Validation of KASP marker Excalibur_c47452_183 (SNP assayed = A/G), closely linked to *Snn3-B1*. Allele-specific primer A: 5′-GAAGGTCGGAGTCAACGGATTaaggctggctggcgagtA-3′. Allele-specific primer B: 5′-GAAGGTGACCAAGTTCATGCTaaggctggctggcgagtG-3′. Common primer: 5′-tgaggggcatcgaatcG-3′. Tails for ligation of fluorophores are included in uppercase at the 5′ end of the allele-specific primers. A subset of 95 variaties from the AM panel were tested.

Table 2.4 – Allele calls for *Snn3-B1* linked SNP marker Excalibur_c47452_183 versus SnTox3 sensitivity score in the AM panel.

	SnTox3 sens. < 4	**SnTox3 sens. = 4**	**Total no. of varieties†**
Allele A:A	317	43	360
Allele G:G	6	91	97
Total no. of varieties†	323	134	457

Sens. = sensitivity.
† Accessions for which phenotypic data were available.

seven SNPs) and the MAGIC population by SMA ($-\log_{10}P > 16$, effect > 1, five SNPs) were aligned to the IWGSC RefSeq assembly (v1.0). The seven SNPs from the AM panel delineated a physical region of 326 kb (6.648 – 6.975 Mb). The five MAGIC SNPs delineated physical regions of 6.2 kb (6.648 – 6.654 Mb) and were all located within the physical interval as defined in the AM panel. The 326 kb region was predicted to contain seven gene models, representing three high-confidence and four low-confidence gene models. Two markers lie within a gene model encoding a ubiquitin-conjugating enzyme: Excalibur_c47452_183 is in the 5′ untranslated

region, while BobWhite_c4838_58 is a synonymous SNP located within exon 6. Three SNPs lie within gene model TraesCS5B01G005000, a P-loop containing nucleoside triphosphate hydrolases superfamily protein: GENE-3324_338 is located in intron 2, BS00091518_51 results in a synonymous substitution in exon 3, and BS00091519_51 is predicted to result in a G→D substitution at amino acid residue 675 (G675D), outside of the 50s ribosome-binding GTPase domain. Finally, two SNPs are located in gene model TraesCS5B01G005600, encoding a putative transmembrane protein: BS00064297_51b and BS00064298_51b both represent non-synonymous mutations (L474P and Q480R, respectively), neither of which are predicted to lie within a known protein domain.

Using BLASTn, with the IWGSC RefSeq wheat genome assembly as the reference, six scaffolds from the cv. Cadenza genome assembly (EIv1) were identified as being within the *Snn3-B1* locus. These scaffolds (Triticum_aestivum_Cadenza_EIv1_scaffold_336839_5BS, Triticum_..._336735_5BS, Triticum_..._338092_5BS, Triticum_..._336737_5BS, Triticum_..._336783_5BS, Triticum_..._336759_5BS) were used in conjunction with FgeneSH to predict gene models in the A×C *Snn3-B1* region. Of the 48 genes predicted on scaffold Triticum_..._336737_5BS, which represents the peak region of the QTL, two were found to be NBS-LRRs, of which one is a protein kinase C (PKC)-NBS-LRR (referred to from here as Gene1; Figure 2.5) and the other is a NBS-LRR-Mildew Locus O (Mlo) (referred to from here as Gene2; Figure 2.6). Gene1 is located within the 327 kb *Snn3-B1* peak region, whereas Gene2 is located approximately 100 kb outside of the peak region. Interestingly using BLASTn and the IWGSC RefSeq genome assembly, Gene1 has a strong partial match (i.e. part of the CDS matches exactly) with TraesCS5B02G005700, however TraesCS5B02G005700, which is from the cv. Chinese Spring IWGSC gene annotation, is lacking LRRs. It is important to note here that cv. Chinese Spring is insensitive to SnTox3 in contrast to cv. Cadenza. Gene2 however is only a 90.7 % match to TraesCS5B02G005700. Both Gene1 and Gene2 are observed to be paralogues of *Tsn1* with BLASTn identity percentages of 87.7 (e-value = $9.4E^{-103}$) and 85.4 (e-value = $3.9E^{-55}$), respectively.

Figure 2.5 – *Snn3-B1* candidate Gene1. The CDS sequence showing protein domains of a PKC-NBS-LRR identified in the *Snn3-B1* locus via *de novo* prediction in SnTox3 sensitive cv. Cadenza. Orange = PKC, blue = NBS, green = LRR.

As one of the most significant *Snn3-B1* marker identified in the MAGIC and AM analyses (GENE-3324_338) was not a perfect predictor of phenotype in the AM panel, it was hypothesised that either (1) genetic recombination was present in the

Figure 2.6 – *Snn3-B1* candidate Gene2. The CDS sequence showing protein domains of a NBS-LRR gene model identified in the *Snn3-B1* locus via *de novo* prediction in SnTox sensitive cv. Cadenza. Blue = NBS, green = LRR, purple = reverse transcriptase, red = Mlo.

AM panel between the SNP and the gene underlying *Snn3-B1*, (2) SNPs in phase with the *Snn3-B1* allele(s) were not present in the iSelect 90,000 feature array used for genotyping, or (3) haplotype analysis would need to be used to increase prediction accuracy. To investigate these hypotheses, genotypic state for the two candidate genes (Gene1 and Gene2) were assessed using a presence/absence PCR/agarose gel assay. This found while the SNP alone had a false positive rate of 25.2 % for allele A (but only 1.2 % for allele G), these false positives could be isolated to a smaller subset of the varieties by combining the SNP allele with the presence or absence of Gene1 and Gene2 (Table 2.5). The rate of false positives improves in groups SNP allele A with Gene1 and Gene2 absent; SNP allele A with Gene1 only present. The majority of false positive cases have been isolated to SNP allele A with both Gene1 and Gene2 present, a group which represents only 17.2 % of the AM panel. It is noteworthy however that, even in combination, none of these factors show complete clustering with the phenotype (Figure 2.7). Therefore, neither the most significant SNP identified by MAGIC and AM panel, nor the candidates genes, provide the basis for a perfect marker for SnTox3 sensitivity alone. Haplotypes constructed by combining the Gene1 and Gene2 genotypic data with that of SNP GENE-3324_338, and comparing these to the SnTox3 phenotypic data in the AM panel found that use of haplotypes improved prediction of the phenotype beyond that of using GENE-3324_338, Gene1 or Gene2 genotypic data alone.

2.5 Discussion

2.5.1 Effector Sensitivity in European Wheat Germplasm and its Relevance to SNB

Septoria nodorum blotch is a major disease of wheat in many growing areas, with field resistance based on multiple minor effect genes. Identification of necrotrophic effectors in *P. nodorum* provided resources with which to dissect host resistance into its constituent parts and study their interactions (Cockram et al., 2015; Phan et al., 2016). Here effector screening was used to determine sensitivities in 480

Table 2.5 – The rate of false positives using SNP GENE-3324_338 and presence of Gene1 and Gene2 to predict SnTox3 sensitivity alone and in combination.

GENE-3324_338	Gene1	Gene2	Score < 3	Score ≥ 3	FPR
A			252	85	25.2 %
G			1	82	1.2 %
	-	-	40	8	16.7 %
	+	-	138	90	39.5 %
	-	+	2	3	40.0 %
	+	+	24	38	38.7 %
A	-	-	36	4	10.0 %
A	+	-	137	31	18.5 %
A	-	+	2	2	50.0 %
A	+	+	24	36	60.0 %
G	-	-	0	4	0.0 %
G	+	-	1	59	1.7 %
G	-	+	0	1	0.0 %
G	+	+	0	2	0.0 %

Initially showing the percentage of false positives for the SNP GENE-3324_338 alone then showing the percentage of false positives using the presence and absence of the candidate genes (Gene1 and Gene2). Finally showing the SNP combined with the presence or absence of Gene1 and Gene2. + = Present. - = Absent. Score = SnTox3 sensitivity score. FPR = false positive rate.

predominantly British winter wheat varieties. The frequency of SnTox1 sensitive varieties was 28 %, broadly comparable to that found for Scandinavian varieties (12 %, Ruud et al., 2018) and global collections (16 %, Shi et al., 2016b), but contrasts notably against a recent screen of Australian varieties (72 %, Tan et al., 2014). SnTox3 sensitivity frequency here (42 %) was similar to that reported in Scandinavian germplasm (55 %, Ruud et al., 2018). Sensitivity to SnToxA was 10 % in the predominantly British winter wheat germplasm collection screened here, which is notably lower than that reported in other wheat germplasm collections, e.g., 45 % in Scandinavian varieties (Ruud et al., 2018) and 65 % in Western Australian wheat (Waters et al., 2011). SnToxA sensitivity was found to be present at a relatively high frequency in the spring wheat varieties in our panel. The spring and winter wheat breeding pools are relatively separate, possibly explaining the observed frequency differences in SnToxA sensitive and insensitive alleles between the two groups. *Tsn1* is located on the long arm of chromosome 5B, < 30 Mb from the vernalization gene *VRN-B1*, a major gene influencing winter or spring growth habit (Cockram et al., 2007). Therefore, it is likely that the partitioning of *Tsn1* alleles is influenced by linkage to winter or spring alleles at *VRN-B1*.

2.5.2 Genetic Mapping of SnTox3 Sensitivity

Previous studies have mapped *Snn3-B1* to the short arm of chromosome 5B (Friesen et al., 2008; Phan et al., 2016; Shi et al., 2016a; Ruud et al., 2017). Here, genetic analysis in the AM, A×C and MAGIC populations identified *Snn3-B1* as representing

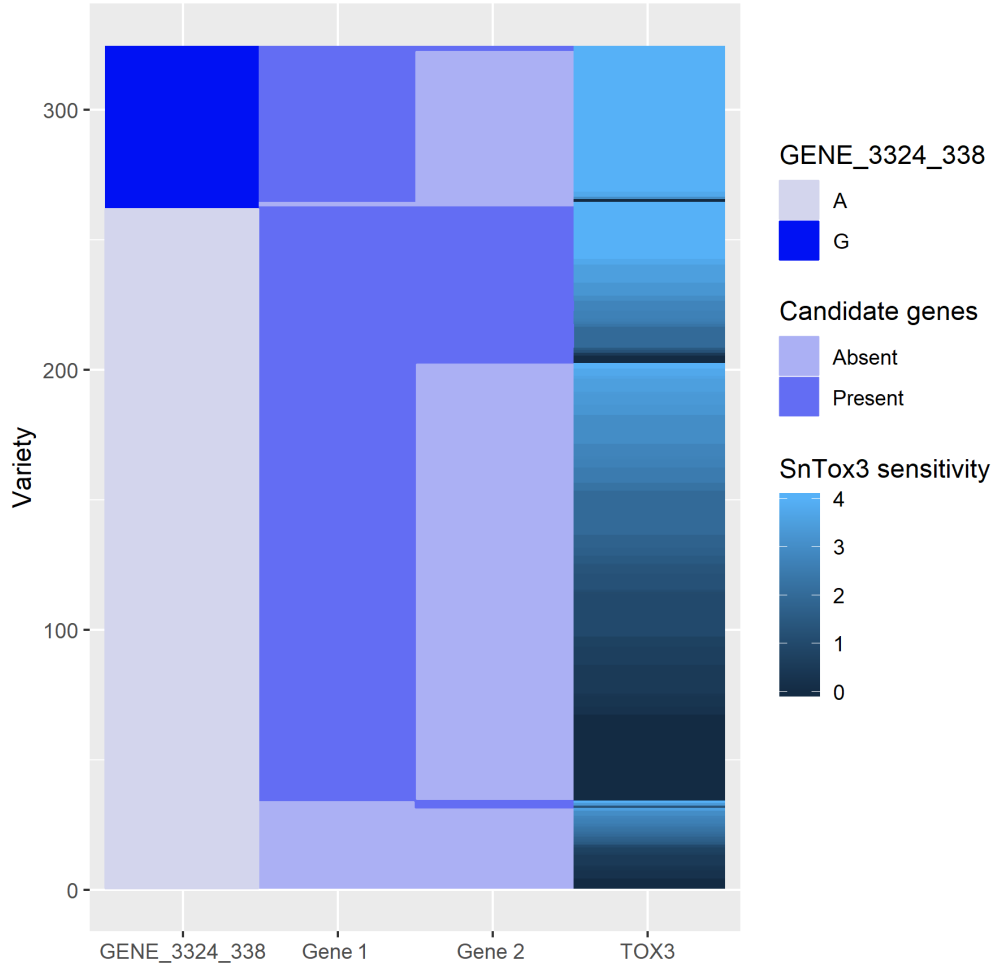


Figure 2.7 – Comparison of the allelic state of genetic markers at the *Snn3-B1* locus with SnTox3 phenotype in the AM panel). Shown are the genotype calls (n = 348) for the most significant iSelect SNP identified in the AM panel and MAGIC population (GENE-3324_338) and the two NBS-LRR candidate genes (Gene1 and Gene2).

the most significant genetic determinant of SnTox3 sensitivity in the European wheat germplasm surveyed, with the most significant markers ($-\log_{10}P > 10$) from the AM panel delimiting a 327 kb interval predicted to contain seven genes in the SnTox3 insensitive variety Chinese Spring. This was achieved without the need to develop populations specifically to investigate SnTox3 sensitivity, demonstrating the efficacy of using genetic resources such as AM panels and MAGIC populations for rapid genetic dissection of target traits (Cockram and Mackay, 2018). The relatively small genetic interval determined here will allow reverse genetic approaches such as genome editing and Targeting Induced Local Lesions IN Genomes (TILLING; McCallum et al., 2000) to be undertaken to help identify the gene underlying *Snn3-B1*. However, it should be noted that given the cultivar Chinese Spring from which the IWGSC RefSeq assembly is derived was found to be insensitive to SnTox3, it is possible that the gene underlying *Snn3-B1* is deleted or has accumulated mutations to such

an extent that it is not predicted as a gene model. This makes Gene1 and Gene2 particularly interesting candidates as they were not present — at least not in their complete form — in the Chinese Spring genome. Indeed, this was the case for the gene underlying the SnToxA sensitivity locus *Tsn1*, which was found to be absent in insensitive varieties, including Chinese Spring (Faris et al., 2010). These examples help illustrate why the availability of genome assemblies for additional wheat varieties, preferably with high quality gene annotations, are potentially so useful for wheat research.

Minor QTL for SnTox3 sensitivity have only previously been reported on chromosome 4B, using a bi-parental population (Phan et al., 2016). The eight minor QTL identified here in the A×C and MAGIC populations, on chromosomes 1B, 2A, 2B, 3B, 4D, 6A, and 7B, are therefore novel. When these are compared with known QTL for SNB field resistance and *P. nodorum* juvenile resistance studies, *QTox3.niab-2A.1* is found to overlap with the SNB resistance QTL *QSnb.niab-2A.4* identified in the same MAGIC population in a multi-year trial in Norway, and located at 758 Mbp on chromosome 2A (**Lin2019**). This QTL also overlaps with another previously identified QTL *Qsnb.cur-2AS.1*, controlling seedling SNB sensitivity using *P. nodorum* isolate SN15, as well as knock-out strains of this isolate lacking SnTox1 (tox1–6), a triple knock-out strain lacking SnToxA, SnTox1, and SnTox3 (toxa13), and seedling inoculation with culture filtrate from isolate toxa13 (Phan et al., 2016). However, as the interval for *QSnb.cur-2AS.1* is very large (112–709 Mbp, based on markers gwm339 and gwm312, respectively), further comparison is not possible. Similarly, the SnTox3 sensitivity QTL identified here on chromosome 3B (*QTox3.niab-3B.1*) co-locate with a QTL for adult plant SNB resistance in the SHA3/CBRD × Naxos wheat population grown under field conditions in Norway, as a number of common markers were identified (Ruud et al., 2017). The creation of near isogenic lines (NILs) for such minor SnTox3 QTL would allow further characterisation of their effects, further investigation of possible impact on Stagonospora Nodorum Blotch (SNB) resistance, and potentially, isolation of their underlying genes. The MAGIC population consists of inbred lines genotyped at the F_5 generation, with each line expected to contain approximately 2 % heterozygosity across the genome, allowing development of heterogeneous inbred families (HIFs) to rapidly create NILs for specific chromosomal locations through selfing (Tuinstra et al., 1997). As MAGIC F_5 lines which are heterozygous across each of the minor QTL are available, it should now be possible to rapidly create precise genetic materials with which to investigate their effects in isolation. It should be noted that while MAGIC QTL analysis using IM/CIM allowed minor QTL to be detected, in comparison to MAGIC SMA (and GWAS analysis in the AM panel), it did not accurately locate *Snn3-B1*. This is due to a 5BS/7BS

chromosomal translocation that is known to segregate in the MAGIC population, with the resulting segregation distortion preventing genetic mapping of markers close to the translocation breakpoint (Badaeva et al., 2007; Gardner et al., 2016). As *Snn3-B1* is close to this breakpoint, the absence of the most closely linked SNPs to the *Snn3-B1* locus in the genetic map prevents accurate mapping via IM/CIM. In contrast, SMA in the MAGIC population does not require markers to be genetically mapped, highlighting the importance of using both analysis methods when undertaking genetic analysis. Indeed, in addition to *Snn3-B1*, SMA and IM/CIM both identified additional QTL, two of which were shared and two and three of which were individual to each analysis method, respectively.

2.5.3 Analysis of the *Snn3-B1* Physical Region

The physical region, as defined by the most significant SNPs identified in the AM panel and MAGIC population, was predicted to contain seven gene models. Gene model TraesCS5B01G005000 (containing SNPs GENE-3324_338, BS00091518_51, and BS00091519_51) is similar to *YELLOW LEAF 1/BRASSINAZOLE INSENSITIVE PALE GREEN 2 (BPG2)*, involved in the accumulation of chloroplast proteins and the salt stress response pathway in Arabidopsis (Li et al., 2016; Qi et al., 2016). The SNP that had the largest phenotypic effect from the MAGIC SMA was located within gene model TraesCS5B01G005100, which encodes a 'ubiquitin-conjugating enzyme E2.' This class of genes has been shown to regulate plant disease resistance, both positively and negatively. Examples include the U-box type E3 ubiquitin ligase, CMPG1, that regulates immunity in multiple plant species, SPL11, a negative regulator of cell death in rice, and Plant U-box 22 (PUB22), PUB23, and PUB24, that negatively regulate PTI in Arabidopsis (Zeng et al., 2004; González-Lamothe et al., 2006; Trujillo et al., 2008). Gene models TraesCS5B01G005200, TraesCS5B01G005300, and TraesCS5B01G005400 all showed sequence similarity to protein kinases, a class of genes known to play a role in disease resistance (Xia, 2004). However, protein kinase domains were only predicted within the amino acid sequence of TraesCS5B01G005400. Finally, gene models TraesCS5B01G005500 and TraesCS5B01G005600 (containing SNPs BS00064297_51b and BS00064298_51b) both encode predicted transmembrane proteins, with BLASTn matches ($\geq 7e^{-66}$) to single, unannotated genes in rice and brachypodium. Further work is needed to investigate whether any of these genes underlie *Snn3-B1*.

The two NBS-LRRs identified from the Cadenza *de novo* annotation of the *Snn3-B1* QTL region, have similar protein domains to *Tsn1*: NBS and LRR. In particular, Gene1 also has a PKC domain in common with *Tsn1*. Indeed PKC-NBS-LRR are particularly rare and *Tsn1* was the first NBS-LRR to characterised with the unique

combination of an N-terminal S/TPK and a C-terminal NBS-LRR (Pandelova et al., 2012). Gene1 and Gene2 CDS were 87.7% and 85.4% similar to *Tsn1*, respectively. The top hit for both predicted genes was TraesCS5B02G005700 with 100% (for part of the sequence) and 90.7% similarity respectively. With both Gene1 and Gene2 from Cadenza having the same top BLASTn hit in Chinese Spring (IWGSC RefSeq v1.0), this raises the prospect of gene duplication with subsequent mutation and functional divergence. Duplication in this manner would allow for neofunctionalisation and can provide an explanation for duplicate retention. Given that Gene1 and Gene2 are the close wheat paralogues to the SnToxA sensitivity gene, *Tsn1*, they appear to be strong *Snn3-B1* candidate genes.

However, when Gene1 and Gene2 were further explored using the presence/absence PCR assay, the resulting data indicated that they may not underlie *Snn3-B1*: while combining the Gene1 and Gene2 presence/absence calls with the allelic state at SNP GENE-3324_338 allowed the prediction of SnTox3 sensitivity to be improved relative to the use of the SNP alone, it did not fully explain the SnTox3 sensitivities observed in the AM panel (Figure 2.7). While the presence or absence of these two genes did not provide a perfect marker for SnTox3 sensitivity, considering their genomic location and similarity to *Tsn1*, they should nevertheless not be excluded as candidate genes at this stage without further evidence.

2.5.4 The Use of Effector Sensitivity Loci for Wheat Research and Breeding

SnTox3 sensitivity and SNB disease susceptibility had previously been reported to be poorly correlated, only accounting for a significant portion of disease phenotype in adult plants segregating for sensitivity alleles at *Snn3-B1*, when infected with *P. nodorum* isolates lacking SnToxA (Friesen et al., 2008). This is consistent with the notion that the SnToxA–*Tsn1* interaction is epistatic to the SnTox3–*Snn3-B1* interaction (Friesen et al., 2008). However, more recently, the *Snn3-B1* locus has been identified in QTL analysis of adult plant SNB resistance in northwestern Europe (Ruud et al., 2017). Furthermore, It is thought that *SnTox1* expression inhibits the transcription of *SnTox3* (Phan et al., 2016). This may explain the reason that while gene-for-gene interactions are readily identified via effector infiltration, their interactions are not always additive. Although, recent work has found infiltration of wheat seedlings with culture filtrate using *P. nodorum* isolates carrying *SnTox3* resulted in a necrotic phenotype on wheat containing *Snn3-B1*, irrespective of the presence of SnTox1 in the pathogen (Ruud et al., 2017). Nevertheless, the differing associations between disease susceptibility and effector sensitivity will likely depend on the effectors present in regional pathogen populations, the interactions between

these effectors, and the alleles present at their corresponding host sensitivity loci. To allow rapid selection for allelic state at *Snn3-B1*, a co-dominant KASP genetic marker closely linked to the locus has been developed in this chapter, for use within wheat breeding programs. The marker represents a useful tool for marker assisted selection for SNB resistance, given the proven association between *Snn3-B1* and SNB resistance in specific agricultural environments. However, despite the relatively simple Mendelian control of the trait in bread wheat, this marker is not a perfect predictor of SnTox3 sensitivity in the 457 phenotyped accessions in the AM panel – most notably, the 43 highly SnTox3 sensitive varieties that carry A:A alleles at SNP Excalibur_c47452_183 (Table 2.4). This is similar to the results of other studies that have attempted to identify diagnostic markers for SnTox3 sensitivity (Shi et al., 2016a; Phan et al., 2018). This observation could be due to a number of reasons, including one or a combination of the following: insufficient marker saturation, multiple alleles at the *Snn3-B1* locus, control by copy number variation, or the effect of minor QTL. Indeed, SnTox3 sensitivity in the AM panel shows more of a quantitative distribution, in contrast to the qualitative phenotypic distribution found for SnToxA sensitivity (Figure 2.2). The observation that combining allele calls from the most significant marker with those from Gene1/Gene2 presence/absence assay indicates that either more than one susceptibility allele is needed, or that other factors such as copy number variation or epigenetic variation may play a role. Nevertheless, the KASP marker and presence/absence markers developed here will be of use in tracking SnTox3 sensitivity alleles, especially where the sensitivity of the founders is known, and to help further narrow the *Snn3-B1* genetic interval. For example, given that all varieties phenotyped that carry G:G alleles, with the exception of two, are highly sensitive (score > 3) for SnTox3, the most significant marker alone could be used to remove the majority of highly sensitive varieties from a breeding program.

3

Transcriptome Investigation of the SnTox3 Response in Wheat

3.1 Abstract

In this chapter RNA-seq and genomic analyses comes together to provide insights into the SnTox3-*Snn3-B1* pathway in wheat. An RNA-seq time course experiment was undertaken over a 24 hour period using two varieties (Avalon, SnTox3 insensitive; Cadenza, SnTox3 sensitive), two treatments (+/- SnTox3 infiltration), five timepoints (immediately after infiltration, and 4, 8, 12 and 24 hours post-infiltration (hpi)) with the aim of identifying differentially expressed genes (DEGs) across the genome as well as candidate genes within the *Snn3-B1* region. After RNA-seq data quality control, 1,837 million paired-end 150 bp Illumina reads aligned to the Cadenza genome containing a pseudomolecule of contigs across the *Snn3-B1* region. Of the 262 genes predicted bioinformatically in the pseudomolecule, 40 were supported by the RNA-seq data generated. Combining the RNA-seq and genomic data, two putative *Snn3-B1* candidate genes were identified. The first, termed gene188 (g188) was identified as the only DEG between SnTox3-infiltrated Cadenza and the remaining variety/treatment combinations (Cadenza SnTox3-, Avalon SnTox3+, Avalon SnTox3-; $p < 0.0001$) within the *Snn3-B1* region. g188 was not expressed in Avalon in any of the timepoints investigated. In Cadenza, while g188 was not expressed immediately after SnTox3 infiltration, it was highly expressed at 4 hpi, with lower expression levels at all

subsequent timepoints. g188 was predicted to encode a galacturonic acid binding wall associated kinase (WAK) which also contains an epidermal growth factor (EGF) domain and serine/threonine domain towards the C-terminus. A second candidate gene, gene187 (g187), located immediately adjacent to g188, was also identified. The RNA-seq-informed g187 Cadenza gene model contained a serine/threonine protein kinase (S/TPK) domain and a motile sperm domain. While g187 was constitutively expressed in both varieties at all timepoints, analysis of the RNA-seq data found the Avalon g187 allele to contain mutations that result in being predicted as a pseudogene. Analysis of genome assembly datasets from 19 wheat varieties indicated three or more alleles to be present at each of the two candidate genes, finding SnTox3 insensitivity to be associated with absence of g188, while intermediate sensitivity and high sensitivity to be associated with two additional g188 alleles. While g187 alleles did not show perfect association with SnTox3 sensitivity, the distribution of g187 alleles across the 19 varieties indicated it is possible they could play a role in mediating SnTox3 sensitivity in varieties carrying functional alleles at g188. Genome-wide, the identification of 32 genes expressed only in Cadenza SnTox3 infiltrated samples, including the rapid upregulation of three ethylene responsive transcription factors, provided key insights into the SnTox3-*Snn3-B1* pathway. Additionally, genes identified as differentially expressed were compared with location of minor SnTox3 sensitivity QTL, while cluster analysis of DEGs helped to further the bioinformatic analysis of key players in SnTox3 response.

3.2 Introduction

Plants undergo attack from a broad range of pathogens (fungal, bacterial, protistic, nematodal, and viral), all of which need to be recognised by the host in order to trigger an immune response. This is achieved through perception of pathogen associated molecular patterns (PAMPs), conserved molecular structures produced by the pathogen (Silva-Gomes et al., 2016). PAMP perception triggers the principle type of plant immunity via pattern recognition receptors (PRRs). The plant immune responses to PAMPs are described as PAMP-triggered immunity (PTI). However, many fungal pathogens also produce effector proteins, which contribute towards virulence in particular (Thomma et al., 2011). Successful pathogen effectors suppress PTI, initiating effector-triggered susceptibility (ETS). Via an arms race between fungal pathogens and plant hosts, a second type of immunity, effector-triggered immunity (ETI) initiated by effector recognition in the host, has evolved in plants (Figure 3.1; Tsuda and Katagiri, 2010).

The evolution from PTI into ETS and ETI is best explained by the zig-zag model (Figure 3.2). The zig-zag model theorises that first PAMPs are detected to induce

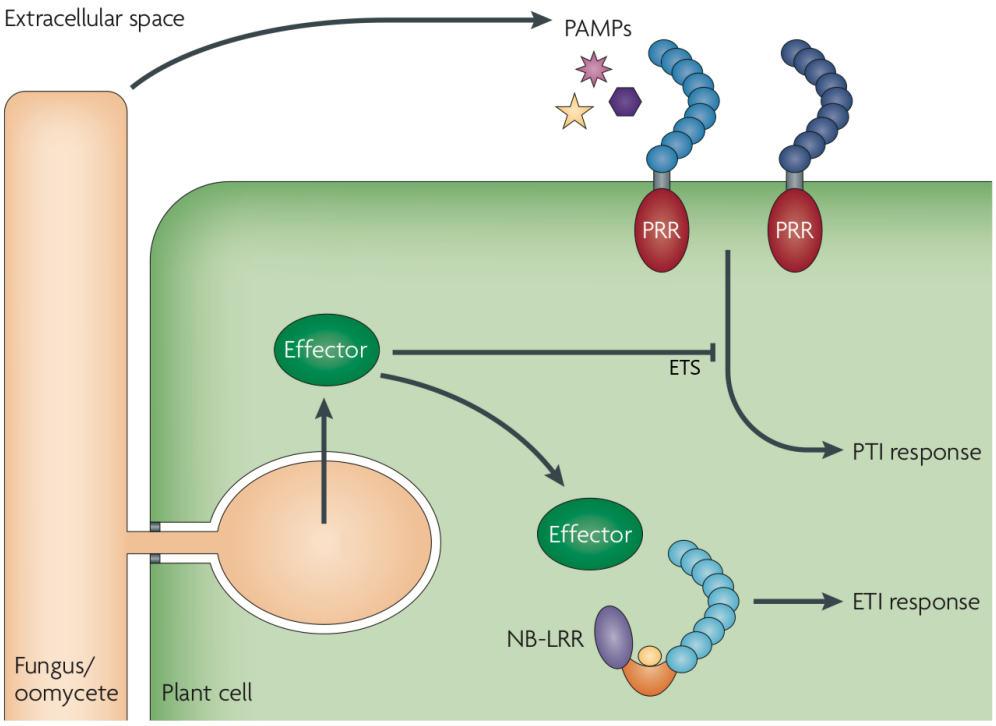


Figure 3.1 – Pathogen associated molecular patterns and effector responses in the host plant. Pathogen associated molecular patterns (PAMPs) are detected by pattern recognition receptors (PRRs), triggering PAMP triggered immunity (PTI). Effectors are recognised by nucleotide binding leucine rich repeats (NB-LRRs) and other Resistance (R) proteins, triggering effector-triggered immunity (ETI). Some effectors can also suppress PTI in the form of effector-triggered susceptibility (ETS). Adapted from Dodds and Rathjen, 2010.

PTI. Then successful pathogens evolve effectors that hijack PTI, enabling pathogen nutrition and dispersal. This results in ETS. However, if then one or a subset of effectors are recognised by a resistance (R) protein, ETI is triggered. The ETI response will often surpass the required threshold for hypersensitive response, which is a stronger immune response than PTI. Therefore pathogens with recognisable effectors have a strong selection pressure against them, in favour of effectors that suppress ETI. In contrast, selection in the host favours individuals with R proteins that can detect these effectors, resulting in ETI (Jones and Dangl, 2006). So ETS can be viewed as a result of hijacking of both the PTI and ETI pathways in the host.

Although it is known that biotrophic fungal pathogens are restricted in growth after PTI and ETI responses, there are a number of known cases of necrotrophs utilising ETS (Dodds and Rathjen, 2010, Shi et al., 2016b). Both PTI and ETI triggers localised programmed cell death. During the early stages of host response to necrotrophic pathogens, analysis of host gene expression datasets show processes linked with host cell death, antimicrobial peptides, ethylene, salicylic acid, abscisic acid, jasmonic acid and the accumulation of reactive oxygen species (ROS) are

initiated (Mengiste, 2012).

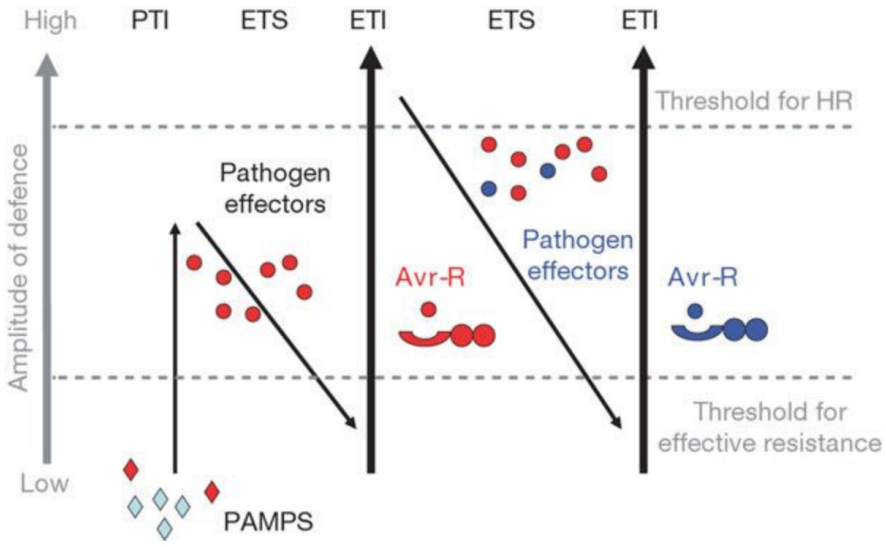


Figure 3.2 – The zig-zag model of plant microbe interactions. Pathogen associated molecular patterns (PAMPs) are detected by PAMP-triggered immunity (PTI), however the effectors of successful pathogens have effectors that trigger effector-triggered susceptibility (ETS). Plants evolved effector-triggered immunity (ETI) to out-compete the pathogen. New effectors, or mutations in existing effectors are able to again, trigger ETS, avoiding host ETI. But then resistant plants in turn evolve recognition proteins for these effectors, again triggering ETI. Taken from Jones and Dangl, 2006.

Known cases of ETS have been observed in cases of oat susceptibility to *Cochliobolus victoriae* (Victoria blight), sorghum to *Periconia circinata* (milo disease), tomato to *Botrytis cinerea* (grey mould), as well as wheat to the necrotrophic pathogens *Pyrenophora tritici-repentis* and *P. nodorum*. The gene *Locus Orchestrating Victorin Effects 1* (*LOV1*), encoding a coiled-coil NBS-LRR, is a member of the *R* gene family. However, *LOV1* has recently been identified as conferring susceptibility to *Cochliobolus victoriae* in *Arabidopsis thaliana* in an inverse gene-for-gene manner. Whereby *LOV1* confers susceptibility to the *C. victoriae* effector victorin, an effector on which *C. victoriae* pathogenicity depends (Lorang et al., 2007). Similarly, the sorghum *Pc* locus has been found to promote sensitivity to a host-selective toxin produced by the fungal pathogen *Periconia circinata* (Nagy and Bennetzen, 2008). Interestingly, *A. thaliana bik1* mutants, which overproduce defense linked hormone salicyclic acid, were found to be more susceptible to the necrotrophic fungal pathogen *B. cinerea*, but were more resistant to the bacterium *Pseudomonas syringae* (Dodds and Rathjen, 2010). Therefore, we can see that ETS to necrotrophic pathogens involves the hijacking of one or more pathogen defense pathways in the plant. The host-pathogen interaction between wheat and *P. nodorum* provides good examples where of the pathogen has hijacked

both the PTI and ETI pathways in the host to induce ETS. The *P. nodorum* effector SnToxA is likely internalised into the plant cell in the leaf before triggering ETI responses (Manning and Ciuffetti, 2005). In contrast, the *P. nodorum* effector SnTox1 interacts with the extracellular component of a wall associated kinase (WAK) encoded by the *Snn1* wheat sensitivity locus, triggering a PTI response (Shi et al., 2016b). Both of these ultimately cause a hypersensitive response in the form of programmed cell death, which is in the interest of the pathogen as it releases nutrient from the cells affected. Thus helping to support continued pathogen infection. Although SnTox3 and *Snn3-B1* were first identified in 2008, the pathways induced and mode of susceptibility conferred by *Snn3-B1* are still relatively unknown (Friesen et al., 2008; Winterberg et al., 2014). This is, in part, due to the unknown identity of *Snn3-B1*. Therefore gene expression datasets from SnTox3 sensitive and insensitive varieties would provide insight into the investigation of the *Snn3-B1* locus and in determining the associated effector recognition pathways involved. This approach is further facilitated by the recent release of the wheat IWGSC RefSeq genome assembly for cv. Chinese Spring (Appels et al., 2018), assemblies of additional wheat cultivars such as Cadenza (EIv1; earlham.ac.uk/grassroots-genomics), and decreasing DNA sequencing costs. The investigation of global changes in gene expression in the cultivars Avalon and Cadenza (insensitive and sensitive to SnTox3, respectively) allows:

1. Expressed genes to be compared to the locations of SnTox3 sensitivity QTLs identified in the Avalon × Cadenza (A×C) doubled haploid population screened in Chapter 2.

2. Exploitation of the recently released Cadenza genome assembly.

Accordingly, analysis of RNA-seq datasets in Avalon and Cadenza with and without SnTox3 infiltration would allow identification of the gene networks involved in response to SnTox3, and help annotate and identify candidate genes within genomic regions underlying wheat quantitative trait locus (QTL) for SnTox3 sensitivity.

3.2.1 Aims

An RNA-seq experiment based on 40 RNA samples collected from the cultivars Avalon and Cadenza across two treatments (+ and - SnTox3 infiltration) and five timepoints (immediately after infiltration and 4, 8, 12 and 24 hours post infiltration), was designed to investigate gene expression in SnTox3 sensitive and insensitive bread wheat lines to:

1. Identify the genes and gene networks involved in the SnTox3 response pathway at the genome-wide level via the identification, classification and analysis of

differentially expressed genes (DEGs) between genotypes and timepoints.

2. Enable the annotation of gene content within SnTox3 sensitivity QTL identified in Chapter 2, and the identification of candidate genes.

3.3 Materials and Methods

3.3.1 Germplasm, Infiltration and Tissue Sampling

Seeds, sourced from the John Innes Centre, Norwich, for wheat varieties Avalon (SnTox3 insensitive) and Cadenza (SnTox3 sensitive) were planted in 96-well trays with M3 (medium/fine) compost and grown in a growth chamber at 20 °C/16 °C with 12 hour photoperiod. At the two leaf stage (Zadok's growth stage 12; Zadoks et al., 1974), half of the seedlings from each of the two varieties were leaf infiltrated with SnTox3 two hours after lights on. Briefly, a 1 ml plastic syringe was used to infiltrate approximately 50 µl of SnTox3 into the first leaf, with the extent of the infiltrated region marked with a non-toxic pen. Tissue samples were collected from the infiltrated leaf (and the uninfiltrated controls) at 0, 4, 8, 12 and 24 hours post infiltration and immediately frozen in liquid nitrogen. For this time course expression study, 50 plants were grown per variety, 25 infiltrated and 25 uninfiltrated. Five leaf samples were taken at random for each variety, timepoint and treatment combination, resulting in a total of 100 tissue samples.

3.3.2 RNA Extraction and RNA-seq

An initial batch of samples, including one sample from each of the five timepoints in Cadenza after SnTox3 treatment, and a bulk of five SnTox3 treated leaf samples from Avalon, was sent to Source Biosciences (Nottingham, UK) for extraction of Total RNA (using the RNeasy Mini Kit [Qiagen, Hilden, Germany]), cDNA library preparation (using the TruSeq Stranded Total RNA with Ribo-Zero rRNA Removal Kit [Illumina, San Diego, CA]) and RNA-seq using the HiSeq 2500 Sequencing System (Illumina, San Diego, CA), running six samples on one lane for a 150 bp paired end run.

For the remaining 35 samples, total RNA was extracted using the RNeasy Mini Kit (Qiagen, Hilden, Germany). RNA was treated with TURBO™ DNase (Invitrogen, Carlsbad, CA) and RNA quality assessed using gel electrophoresis and quantified using a Nanodrop 8000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). Oligo(dT) enrichment of mRNA, fragmentation, cDNA synthesis and PCR amplification was carried out by Novogene (Beijing, China), with RNA sequencing performed using a HiSeq 2500 Sequencing System (Illumina, San Diego, CA) for two

replicates of Cadenza at each timepoint under SnTox3 treatment, three replicates of Avalon each timepoint under SnTox3 treatment and one replicate of each control (i.e. without SnTox3 infiltration) for each variety and timepoint in a 150 bp paired end run.

3.3.3 Mapping RNA-seq Reads to the Cadenza Genome Assembly

Cadenza scaffolds (EIv1.0; earlham.ac.uk/grassroots-genomics) were aligned to the bread wheat cv. Chinese Spring RefSeq genome assembly (v1.0; Appels et al., 2018) with MiniMap2 (v2.15; Li et al., 2018) using commandline options: -I 20G -ax asm5 –secondary=no. Scaffolds aligning to the *Snn3-B1* region (interval defined by marker BS00065165 from SHA3/CBRD × Naxos and marker xtc252302 from Calingiri × Wyalkatchem; Ruud et al., 2017; Phan et al., 2016) of Chinese Spring chromosome 5B (4.862–7.541 Mbp) were extracted from the Cadenza assembly and ordered and orientated based on their sam file coordinates and mapping orientation flags before being joined together as one fragment with each contributing scaffold separated by 100 Ns. The assembled fragment (termed as the 'Cadenza *Snn3-B1* pseudomolecule') was also reintroduced to the Cadenza assembly having removed the contributing scaffolds. The Cadenza pseudomolecule was produced by L. Percival-Alwyn.

The RNA-seq reads, produced by Novogene and Source Biosciences using the 40 RNA samples (three biological reps per timepoint per variety for infiltrated samples; one biological rep per timepoint per variety for uninfiltrated samples), were trimmed to remove adaptors and quality filtered using TrimGalore (Krueger, 2012) prior to genome alignment using HISAT2 looking at primary alignments only (Kim et al., 2017).

3.3.4 Quantification

A *de novo* gene model for the Cadenza *Snn3-B1* pseudomolecule was produced using the software AUGUSTUS (Stanke et al., 2008) with species setting 'wheat', all other settings were default. Predicted protein sequences produced by AUGUSTUS were categorised using the NCBI Conserved Domain Database (Marchler-Bauer et al., 2010) and the proteins predicted by AUGUSTUS were labelled in consecutively, i.e. g1, g2.... g262. The Cadenza *Snn3-B1* pseudomolecule predicted gene models and the Chinese Spring IWGSC RefSeq gene models (v1.0; Appels et al., 2018) were used with the RNA-seq reads to calculate gene counts for the pseudomolecule predicted genes as well as the whole Chinese Spring genome using Kallisto (Bray et al., 2016). R/DESeq was then used to normalise the gene counts and analyse for DEGs

(Anders and Huber, 2010). DEGs were identified using a chi-squared test comparing two generalised linear models to specifically identify genes that are significantly differentially expressed between both treatment and variety, i.e. identifying DEGs specific to SnTox3-infiltrated Cadenza relative to the other groups (Avalon SnTox3+, Avalon SnTox-, Cadenza SnTox-; adjusted p-value < 0.05).

3.3.5 DEG Clustering Analysis

principal component analysis (PCA) was carried out on all significantly DEGs that were expressed after infiltration with SnTox3 using the prcomp function in R (R Core Team, 2013). Grouping genes, with similar expression patterns over time in Cadenza samples after SnTox3 infiltration, into six clusters. This was carried out in R/kohonen (v3.0.8; Wehrens and Kruisselbrink, 2018) using self-organising maps and the PCA data, dividing the DEGs into 36 classification units, which then formed the six clusters. The optimal number of clusters was determined using the gap statistic method and k-means. The gene ontology (GO) Slim Terms were identified for the genes within each cluster using the wheat GOMAP annotations (Lawrence-Dill, 2019).

3.3.6 Comparisons of *Snn3-B1* Region Genomic Sequences

Seven of the eight NIAB MAGIC Elite Founder genome assemblies — Alchemy, Brompton, Rialto, Soissons, Xi19 (NIAB/John Innes Centre/ Earlham Institute/ Natural History Museum), Claire and Robigus (earlham.ac.uk/grassroots-genomics) — were mapped to the Cadenza *Snn3-B1* pseudomolecule using minimap2 (Li, 2018). Raw reads underlying these assemblies were also mapped to the Cadenza *Snn3-B1* pseudomolecule using Burrows-Wheeler Aligner (BWA) (Li and Durbin, 2010) looking at primary alignments only. In addition, presence of the candidate genes was ascertained using BLASTn with the following genomes: seven NIAB MAGIC Elite founders (detailed above), Arina, Jagger, Julius, Lancer, Mace, Norin61, SY-Mattis, Zavitan, (10+ Wheat Genomes Project; webblast.ipk-gatersleben.de/wheat_ten_genomes/), Chinese Spring (IWGSC RefSeq v1.1) and Paragon (earlham.ac.uk/grassroots-genomics). Utilising RNA-seq data, protein sequences were ascertained using EMBOSS Transeq (Rice et al., 2000) and protein domains were identified using Interpro (Mitchell et al., 2018).

3.3.7 TILLING

M_5 Targeting Induced Local Lesions In Genomes (TILLING) lines (n = 1100) from cv. Cadenza (Krasileva et al., 2017), sourced from the John Innes Centre (Norwich,

UK) were grown in 96-well trays with fine/medium compost (M3) in a heated and lit glasshouse at 20 °C/17 °C day/night with 16 hour photoperiod. Each line was represented by three replicates. The seedlings were infiltrated as described in section 3.3.1. Seven days following infiltration, the plants were visually evaluated for SnTox3 effector sensitivity on a scale of 0 (insensitivity, no symptoms) to 4 (extensive necrosis; Tan et al., 2012). Lines displaying hyposensitivity to SnTox3 in at least two of three replicates were selected for further work. DNA was extracted from leaf tissue of selected TILLING lines using a crude DNA extraction method: leaf material was disrupted using ballbearings in 400 µl crude DNA extraction buffer (200 mmol · dm^{-3} trizma® base (Sigma-Aldrich, St. Louis, MO); 250 mmol · dm^{-3} sodium chloride (NaCl); 25 mmol · dm^{-3} ethylenediaminetetraacetic acid (EDTA); 0.5 % sodium dodecyl sulfate (SDS); adjusted to pH 7.5, made up to 1 l; filtered and autoclaved); incubated for 1 hour at 65 °C; centrifuged; supernatant transferred to 400 µl isopropanol; centrifuged; before re-suspending the pellet in TE buffer (10 mmol · dm^{-3} Tris-hydrochloride (Tris-HCl); 1 mmol · dm^{-3}).

For target genes which possessed previously published exome capture sequence data (Krasileva et al., 2017), TILLING mutations were identified using the 'EMS Cadenza' track in the Ensembl Plants genome browser (plants.ensemnl.org/Tritcum_aestivuminfo/index; Bolser et al., 2017). For target genes not present in the exome capture array previously and for which no information was available in Ensembl Plants, homoeologue specific primers were designed to amplify the genes from predicted start codon to predicted stop codon in overlapping segments. Primers were ordered from Sigma-Aldrich (St. Louis, MO). PCR reactions were carried out as follows: 5 ng of Cadenza TILLING DNA mixed with 0.2 µl FastStart™ *Taq* DNA polymerase (Roche, Basel, Switzerland), 2.5 µl FastStart ×10 buffer without magnesium chloride ($MgCl_2$), 2.5 µl (25 mmol · dm^{-3}) $MgCl_2$ and 0.5 µl deoxynucleoside triphosphates (dNTPs) (10 µmol · dm^{-3} of each dNTP), 1 µl (10 µmol · dm^{-3}) of each primer (forward and reverse) and made up to 20 µl using polymerase chain reaction (PCR) grade water. PCR reactions were incubated at 96 °C for 60 s followed by 35 cycles of 50 s at 96 °C, 50 s at 60 °C and 180 s at 72 °C, followed by a final incubation at 72 °C for 240 s. All PCRs were carried out using a Veriti 96 well Thermal Cycler (Applied Biosystems, Waltham, MA), followed by agarose gel electrophoresis and gel extraction using the GeneJET Gel Extraction Kit following the manufacturer's instructions (Thermo Scientific, Waltham, MA). Purified PCR amplicons were sent to GeneWiz (Bishop's Stortford, UK) for Sanger sequencing using the primers used for PCR amplification.

3.4 Results

3.4.1 Pseudomoleculing and *De Novo* Gene Prediction of the *Snn3-B1* Genomic Region in Cadenza

To investigate gene expression within the *Snn3-B1* locus, a pseudomolecule of contigs across the *Snn3-B1* region was first created using the publicly available genome assembly of the SnTox3 sensitive variety Cadenza (EIv1). 151 Cadenza contigs (with an average length of 17.2 kb, total length 2.60 Mb) were found to map to the 2.7 Mb *Snn3-B1* region in the Chinese Spring wheat reference genome (IWGSC RefSeq v1.0). 262 genes were predicted by AUGUSTUS to be present in the Cadenza *Snn3-B1* pseudomolecule. Using Interpro to identify protein domains, 26 % of the predicted proteins within the Cadenza *Snn3-B1* pseudomolecule were found to be linked to tranposons: 9.1 % were identified as part of the retrotransposon family, 6.5 % were annotated as transposonases and 10.3 % were identified as part of the reverse transcriptase family. Of the 74 % of gene models not annotated as (or linked to) transposable elements, the following encoded predicted proteins of possible note, based on *a priori* knowledge of plant disease resistance gene families: two WAK domain genes, five protein kinase C (PKC) genes, two leucine-rich repeat (LRR) genes, one NB-ARC domain gene and two serine/threonine kinase - interleukin-1 receptor-associated kinase (STKc-IRAK) genes (Appendix B.1). This annotated Cadenza *Snn3-B1* pseudomolecule served as a genomic reference from a SnTox3 sensitive variety onto which subsequent RNA-seq data derived from Cadenza and Avalon could be superimposed.

3.4.2 Analysis of RNA-seq Data Aligning to the *Snn3-B1* Locus

RNA-seq was undertaken using the varieties Avalon and Cadenza, representing a SnTox3 insensitive genotype and SnTox3 sensitive genotype, respectively. These varieties were infiltrated with SnTox3 and leaf tissue sampled from the infiltrated leaf at 0, 4, 8, 12 and 24 hours after infiltration with three biological replicates for each timepoint. A control group of un-infiltrated plants was also sampled with one biological replicate for each timepoint. cDNA libraries were constructed from second leaves of the seedling, with subsequent RNA-seq generating 1,844 million 150 bp paired-end RNA-seq reads. After trimming and quality control, 1,837,494,233 paired-end reads were aligned to the Cadenza genome, which included the Cadenza 2.6 Mb *Snn3-B1* pseudomolecule. Of these, 170,872 paired-end RNA-seq reads were mapped to the Cadenza *Snn3-B1* pseudomolecule. Of the 262 genes predicted by AUGUSTUS on the Cadenza *Snn3-B1* pseudomolecule, 40 had RNA-seq expression

reads aligning in least one of the 40 Avalon and Cadenza samples sequenced. Within the Cadenza *Snn3-B1* pseudomolecule, only one DEG was identified, gene188 (g188), as significantly differentially expressed between SnTox3-infiltrated Cadenza and the other groups (i.e. Cadenza SnTox3-, Avalon SnTox3+, Avalon SnTox3-; adjusted-P < 0.0001; Figure 3.3). The reason it was identified is while g188 is expressed in Cadenza after SnTox3 infiltration, g188 was not observed in the SnTox3 insensitive line Avalon, nor was it observed in the SnTox3 sensitive line Cadenza in the absence of SnTox3 infiltration. g188 expression was not observed immediately after leaf infiltration with SnTox3 either. However, a spike in g188 expression at 4 hours post-infiltration (hpi) was observed with a mean normalised gene count of 27, which then decreased over the subsequent three timepoints at 8, 12 and 24 hpi. The g188 gene counts at 4, 8 and 12 hpi are statistically significantly different from each other, as determined by t-test ($p < 0.05$). BLASTn searches of the wheat genome using g188 CDS find the equivalent gene in the reference genome of SnTox3 insensitive Chinese Spring to be TraesCS5B02G005755 (e-value = 0, ID = 92 %). TraesCS5B02G005755 is annotated as a 'nontranslating' CDS. Manual translation of TraesCS5B02G005755 CDS results in four predicted stop codons, the first of which is located at the start of the protein kinase domain. Together, this suggests a non-functional allele of g188 is present in the SnTox3 insensitive variety Chinese Spring.

3.4.3 Additional candidate genes in the Cadenza *Snn3-B1* region

Although no other genes were significantly differentially expressed within the Cadenza *Snn3-B1* pseudomolecule, one additional gene was of note: gene187 (g187), located immediately adjacent to g188. g187 was predicted to contain a serine/threonine protein kinase (S/TPK) domain towards the N-terminus and a motile sperm domain towards the C-terminus. While g187 represented one of the 40 bioinformatically identified Cadenza gene models backed-up by RNA-seq, it was found to be constitutively expressed in both Avalon and Cadenza across all timepoints and treatments. However, comparison of RNA-seq data between the two varieties found it to have a 1 base pair (bp) exonic insertion resulting in a frame-shift and a premature stop codon (K153KfsX11) in Avalon and four exonic single nucleotide polymorphisms (SNPs) (two prior to the frameshift: C125C, E152K; Figure 3.4). Additionally, no RNA-seq read coverage was observed at the 5′ (595 bp) and 3′ (115 bp) end of the gene in Avalon. This indicates that the SnTox3 insensitive variety Avalon carries a non-functional allele of g187. g187 and g188 genes fall within the most significant region of the *Snn3-B1* locus at 1,802 - 1,811 kb in the Cadenza *Snn3-B1* pseudomolecule, nevertheless it is clear that this is a

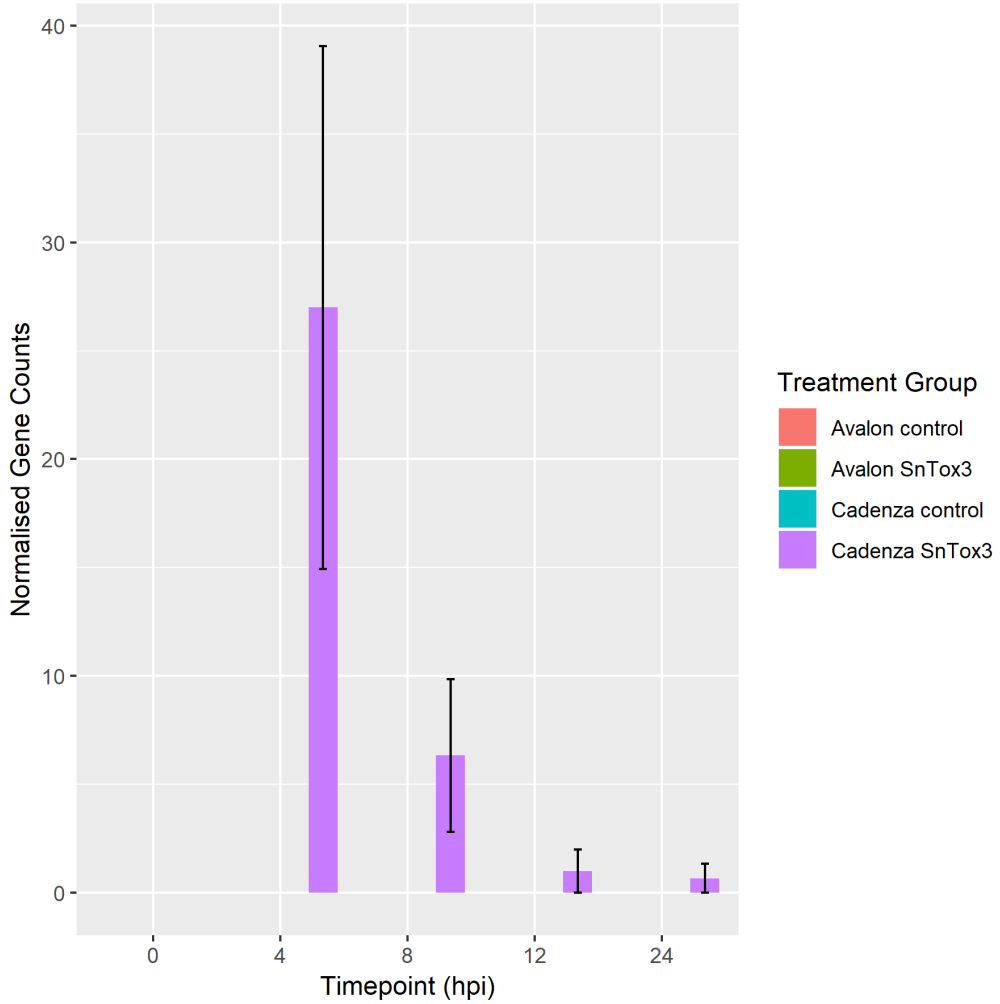


Figure 3.3 – Differential expression of Gene188 (g188), located in the Cadenza *Snn3-B1* genomic region. Expression levels (in normalised gene count) of g188 with an adjusted p-value of < 0.0001 between SnTox3-infiltrated Cadenza and the other samples (Avalon SnTox3+, Avalon SnTox3-, Cadenza SnTox3-), shown across timepoints 0, 4, 8, 12 and 24 hours post infiltration (hpi). g188 gene counts at 4, 8 and 12 hpi are significantly different from each other, as determined by t-test ($p < 0.05$). g188 was not expressed in Avalon or the control samples at any of the timepoints investigated.

complex genomic region. Kmer analysis shows that the level of duplication in this region is quite high (Figure 3.5). No other suitable candidates were identified based on RNA-seq and it should be noted that the two NBS-LRR candidates (Gene1 and Gene2), identified in Chapter 2, were not observed.

3.4.4 Analysis of RNA-seq Data Across the Whole Genome

Statistical analysis using R/DESeq identified 18,563 genes (identified using the Chinese Spring IWGSC high- and low-confidence gene models plus the Cadenza *Snn3-B1* pseudomolecule *de novo* annotation generated in this chapter) as differentially

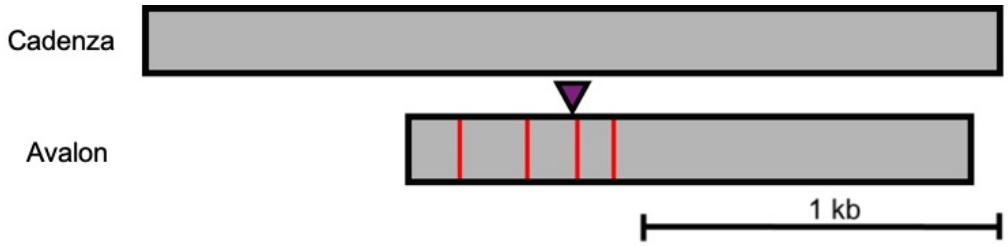


Figure 3.4 – Differences in g187 coding sequences (CDS) predicted from RNA-seq data between the SnTox3 sensitive variety Cadenza and insensitive variety Avalon. Cadenza and Avalon differ by four SNPs and a 1 bp insertion, at 667, 816, 894, 957 and 892 bp, relative to the Cadenza sequence. In addition, no Avalon RNA-seq reads were identified across 595 and 115 bp regions at the 5′ and 3′ ends relative to the Cadenza gene model, respectively.

expressed between SnTox3-infiltrated Cadenza and the other samples (Cadenza SnTox3-, Avalon SnTox3+, Avalon SnTox3-). When analysed by individual timepoint, comparing Cadenza after SnTox3 infiltration with Avalon and control lines, 2446, 766, 1326, 1010, 1921 DEGs were found to be upregulated at 0, 4, 8, 12, 24 hpi, respectively. 674, 554, 1178, 512 and 873 DEGs were found to be down-regulated in response in SnTox3 treatment in SnTox3 sensitive Cadenza at each timepoint, respectively (Figure 3.6).

5 % of these DEGs were categorised via their IWGSC gene model annotations as transcription factors, salicylic acid regulated genes, ROS regulation and signalling genes, receptor protein kinases, genes involved in programmed cell death, nucleotide binding site (NBS)-LRR proteins, jasmonic acid related genes, exo-/endo-cytosis related proteins, ethylene related genes and antifungal proteins. When summarised categorically an overall up-regulation of expression in these pathways over the 24 hours following SnTox3 infiltration in Cadenza is observed (Figure 3.7). 52 transcription factors are up-regulated immediately after infiltration, along with genes involved in ethylene-related processes (n = 49) and jasmonic acid (n = 33). For genes linked with ROS regulation, receptor protein kinases, NBS-LRR, antifungal proteins (including Mitogen-activated protein kinase 3 [MPK3] and pathogenesis-related 1 [PR-1] proteins) and programmed cell death, up- and down-regulation remains fairly consistent in the first 24 hour period, with the exception of 4 hpi where a decrease in up-and down-regulation across all of these pathways is observed.

In addition, it was found that 336 DEGs are expressed only after SnTox3 exposure in sensitive or insensitive lines. 51 DEGs were found to be only expressed in Avalon after SnTox3 treatment, but not in Cadenza. After SnTox3 infiltration 32 DEGs were detected that were expressed only in the Cadenza SnTox3 infiltrated samples (Figure 3.8). These are termed ‘switched-on DEGs’ from here on, and were found to be linked with camalexin biosynthesis, jasmonic acid biosynthesis, abscisic acid

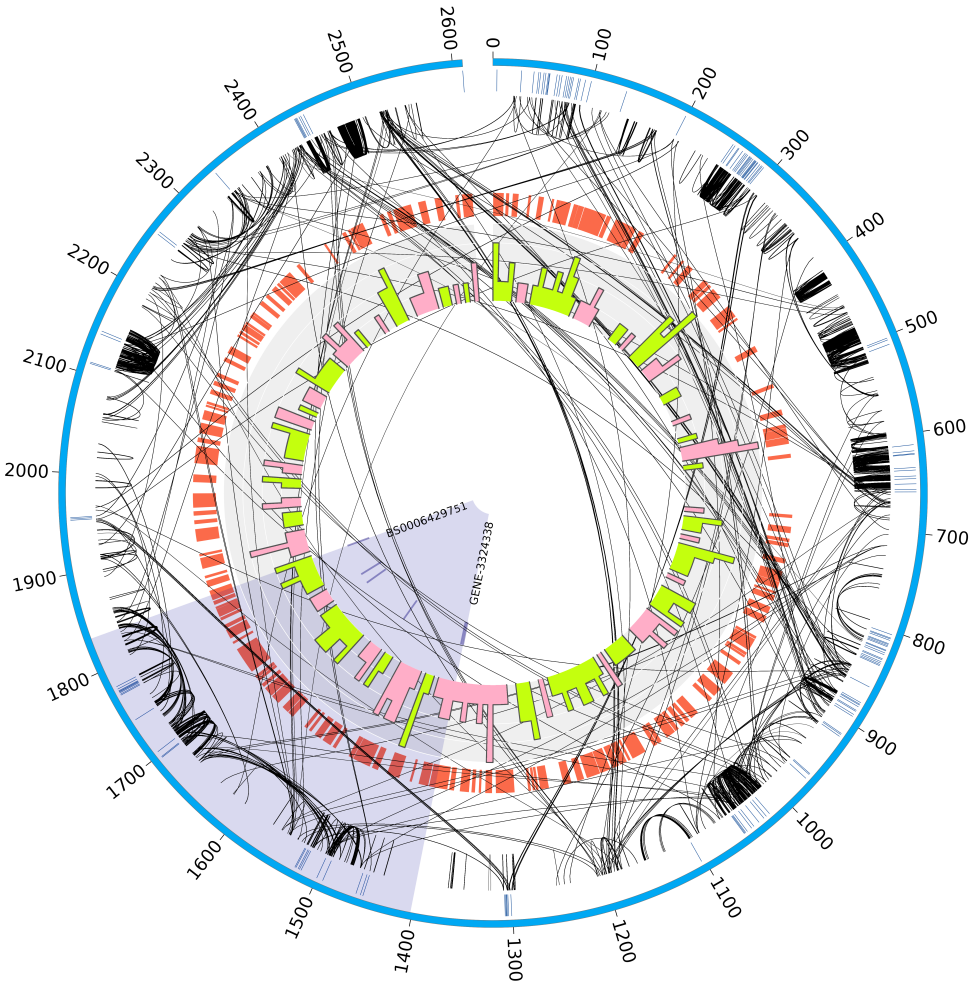


Figure 3.5 – Circos plot illustrating features of the *Snn3-B1* region pseudomolecule constructed for SnTox3 sensitive variety, Cadenza. The outer blue track represents the 2.6 Mb Cadenza pseudomolecule. The dark blue ticks show the joins between Cadenza scaffolds. The links in black show the Kmer duplications. The orange ticks represent genes in the Cadenza AUGUSTUS gene annotation (Chapter 3). The green and pink histogram track shows the frequency of genes present. The highlighted blue region between GENE-3324_338 and BS00064297_51 is the region identified via GWAS in MAGIC and AM (Chapter 2). Displayed in Kb.

biosynthesis, phenylpropanoid biosynthesis, lignin metabolic processes, hydrogen peroxide catabolic processes and trehalose biosynthesis. Moreover, of the 32 'switched on' DEGs, twelve, seven and five were linked to disease resistance, fungal pathogens and cell death, respectively, according to KnetMiner (Hassani-Pak et al., 2016).

Specifically, three ethylene-responsive transcription factors were found to have an extremely high combined average normalised gene count of 1881 immediately after infiltration, but decreasing to an average gene count of one at 4 hpi (Figure 3.8). This pattern of expression is different to DEGs encoding protein kinases, which are seen, for the most part, to steadily increase over the 24 hour period or the SPX (named

after: Syg1, Pho81 and XPR1) domain-containing genes that steadily increase, peak at 8 hpi, and then decrease.

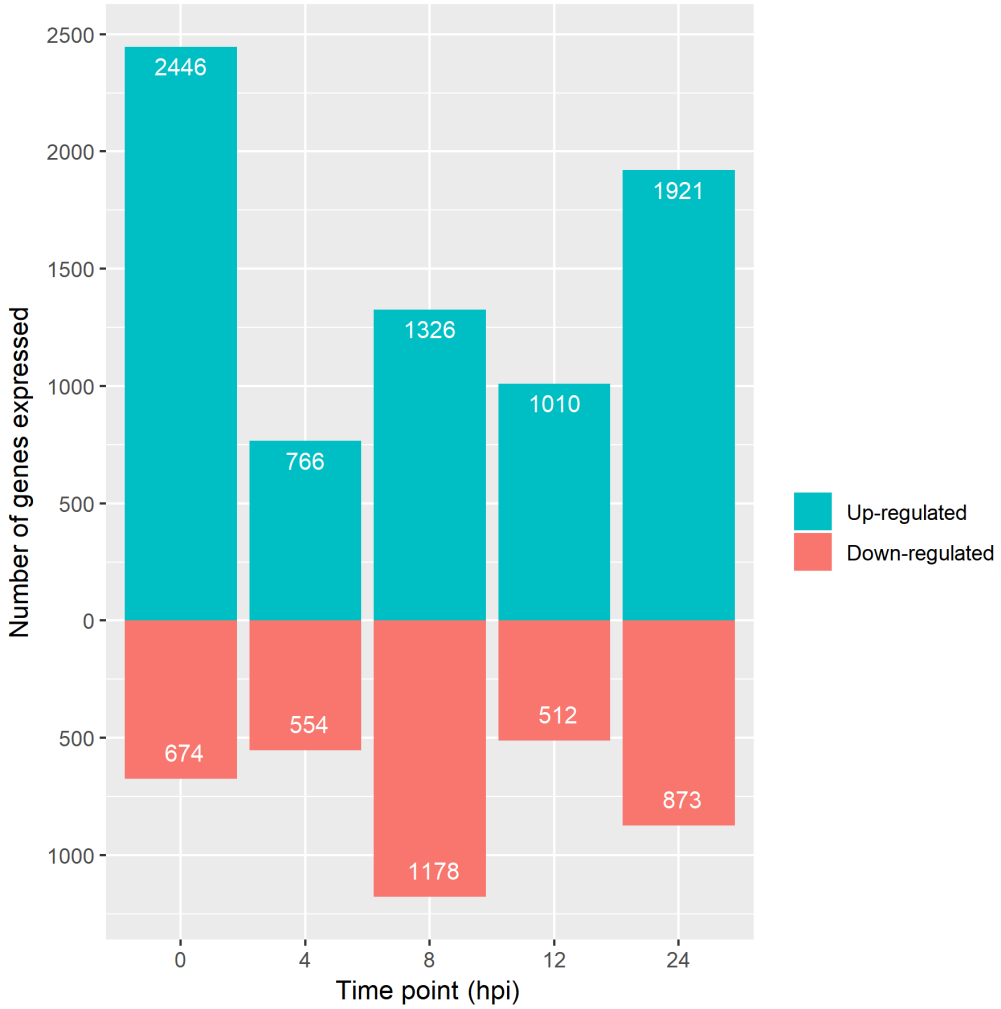


Figure 3.6 – Up- and down-regulated differentially expressed gene (DEGs) over time post-infiltration. The number of DEGs up- or down-regulated at each timepoint in Cadenza under SnTox3 treatment compared with the all other samples.

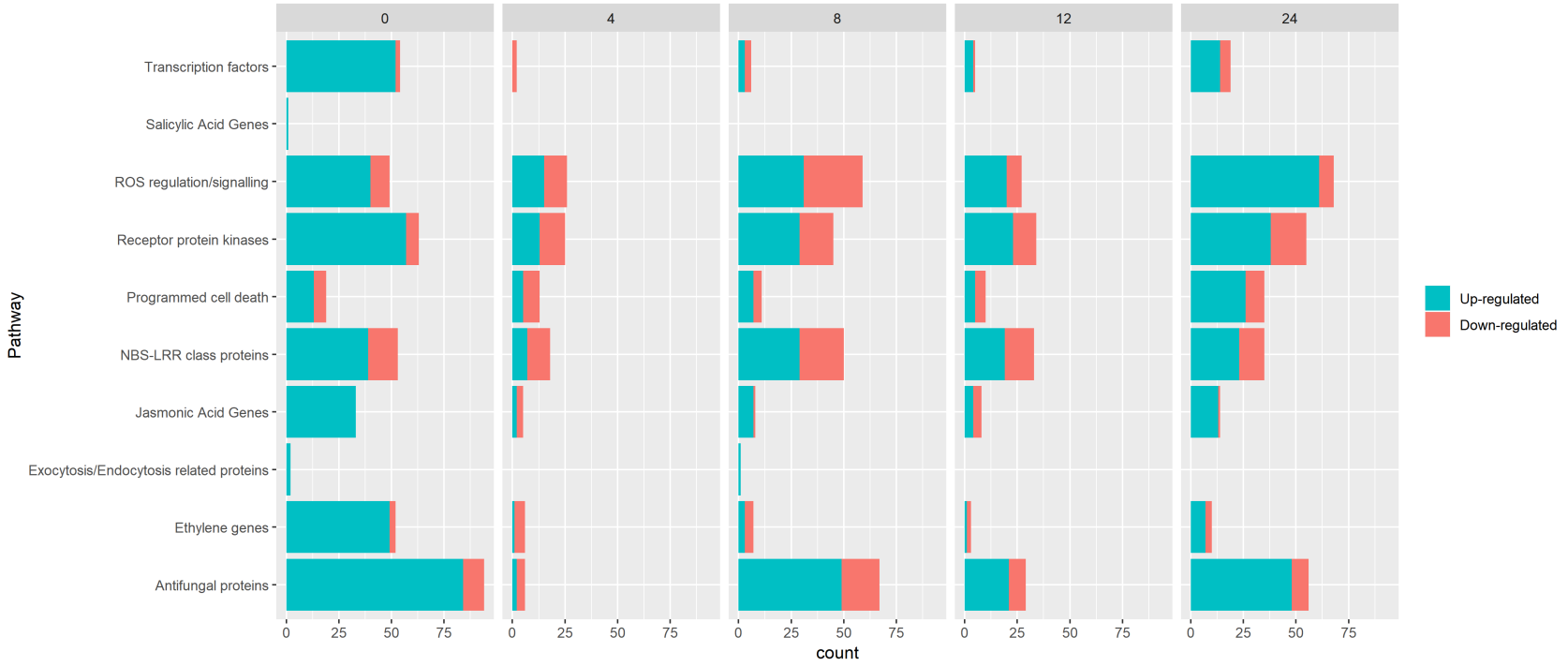


Figure 3.7 – Up- and down-regulation of genes belonging to specific pathways and processes after SnTox3 infiltration. The number of differentially expressed genes up- or down-regulated at each timepoint in Cadenza under SnTox3 treatment compared with all other samples in specific pathways (identified using the IWGSC RefSeq gene annotations). ROS = reactive oxygen species, NBS-LRR = nucleotide binding site - leucine rich repeats.

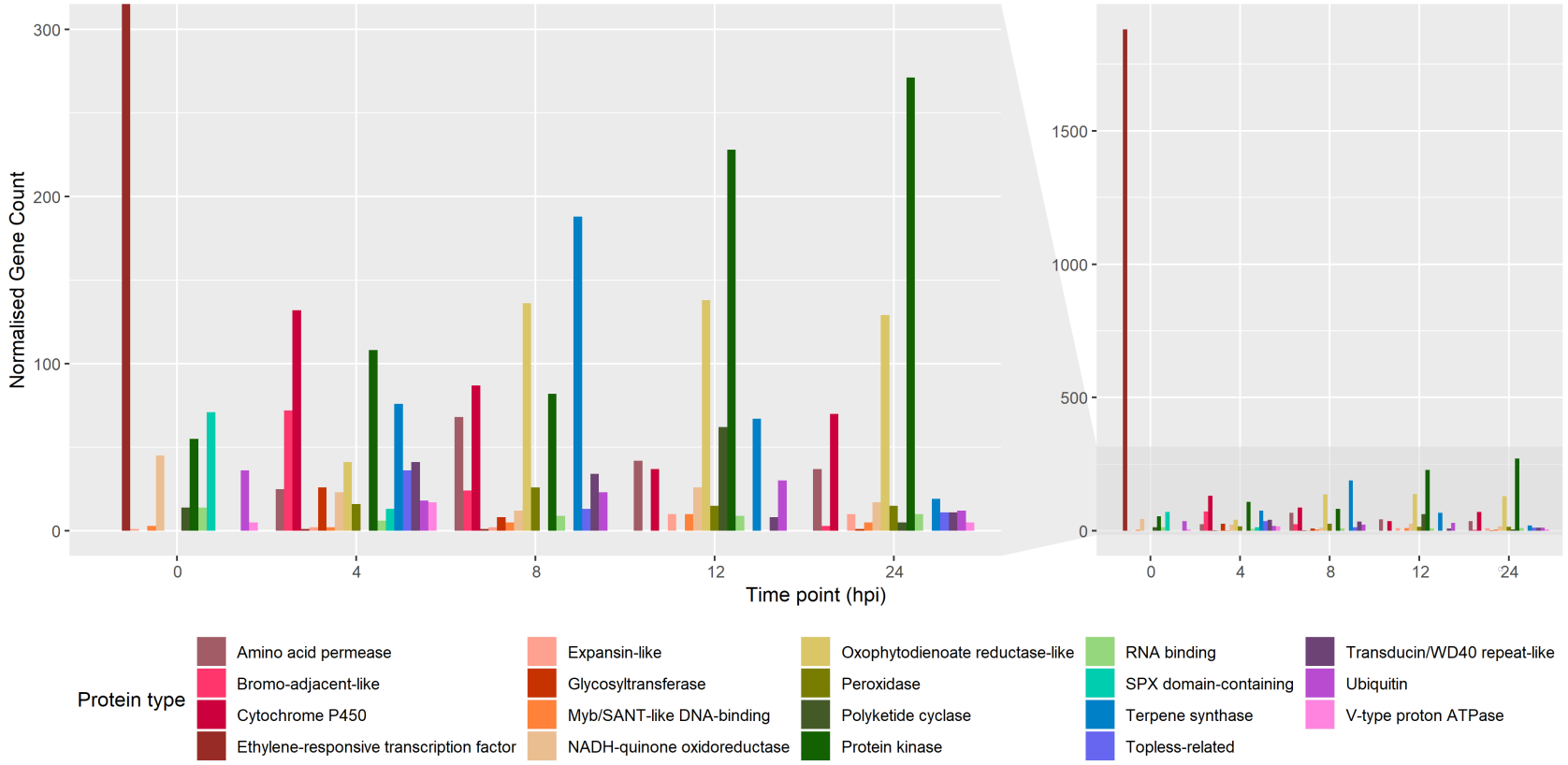


Figure 3.8 – Expression of the 32 genes that are expressed only in SnTox3 infiltrated Cadenza samples (termed 'switched on' genes) over time. These genes fall into 19 gene families with ethylene-responsive transcription factors showing an extremely notable peak in gene expression immediately after infiltration. The histogram on the left is a zoomed in subsection of that on the right. Hpi = hours post SnTox3 infiltration.

3.4.5 DEGs Co-locating to Minor SnTox3 Sensitivity QTL

Of the nine minor bread wheat SnTox3 response QTL, located on chromosomes 1B, 2A, 2B, 3B, 4D, 6A, 6B, 7B and defined with the flanking markers (Chapter 2; Table 3.1), genes were significantly differentially expressed ($P < 0.05$) within three of these: *QTox3.niab-2B.1* (NIAB MAGIC Elite), *QTox3.niab-3B.1* (NIAB MAGIC Elite), *QTox3.niab-7B* (NIAB MAGIC Elite; Table 3.1). Four genes within *QTox3.niab-2B.1* are expressed after exposure to SnTox3 in Avalon only: G-box regulating factor 6, a receptor-like protein kinase, acyl protein thioesterase 1 and a peroxidase. One DEG was identified within the *QTox3.niab-3B.1* interval, an amino acid permease, and seven genes within *QTox3.niab-7B.1*: a protein detoxification gene, pheophorbide a oxygenase, a kinase, Ferredoxin NADP reductase 2, a dirigient protein, a RING finger protein and a zinc finger protein (Table 3.1). Previously cloned *P. nodorum* susceptibility genes, *Tsn1* (conferring sensitivity to SnToxA) and *Snn1* (SnTox1) were expressed in both Avalon and Cadenza, but were not differentially expressed.

Table 3.1 – *P. nodorum* SnTox3 QTL

QTL	Location (Chr, Mb)	Peak markers†	Reference	DEGs
Snn3-B1	5B, 6.654, 6.647	Excalibur_c47452_183, GENE-3324_338	R. C. Downie et al., 2018	Gene188
QTox3.niab-1B.1	1B, 349.715	Tdurum_contig76507_422	Chapter 2	
QTox3.niab-2A.1	2A, 758.692, 758.554	BS00070979_51, Excalibur_c20478_641	R. C. Downie et al., 2018	
QTox3.niab-2B.1	2B, 799.254	Kukri_c9898_1766	R. C. Downie et al., 2018	TraesCS2B02G519700, TraesCS2B02G563900, TraesCS2B01G573500, TraesCS2B02G614100
QTox3.niab-2B.2	2B, 771.205	Excalibur_c30571_95	Chapter 2	
QTox3.niab-3B.1	3B, 67.942	wsnp_Ex_c11246_18191331	R. C. Downie et al., 2018	TraesCS3B02G441300
QSnb.cur-4BL	4B, 641.521	wmc413, wpt-730303	Phan et al., 2016	
QTox3.niab-4D.1	4D, 32.347	BS00036421_51	R. C. Downie et al., 2018	
Snn3-D1†	5D, 5.597	xcfd18	Zhang et al., 2011	
QTox3.niab-6A.1	6A, 22.443, 22.524	BobWhite_c13839_135, IACX7801	R. C. Downie et al., 2018	
QTox3.niab-6B.1	6B, 8.410	wsnp_Ku_c2119_4098330	R. C. Downie et al., 2018	
QTox3.niab-7B.1	7B, 3.541	BS00022127_51	R. C. Downie et al., 2018	TraesCS7B02G032300, TraesCS7B02G038100, TraesCS7B02G041100, TraesCS7B02G045300, TraesCS7B02G050800, TraesCS7B02G061100, TraesCS7B02G065700

Summary of published wheat sensitivity QTL for *P. nodorum* effector SnTox3, showing the DEGs found in each interval.
† This QTL was found in *Aegilops tauschii* rather than *Triticum aestivum*.

3.4.6 Clustering Analysis of DEGs

PCA of the 18,563 DEGs showed that 83 % of the variation between expression patterns of SnTox3-infiltrated Cadenza plants over time were represented in three principal components (Figure 3.9). Six clusters of gene expression patterns were determined (Figure 3.10). All clusters contained genes involved in carbohydrate metabolism, lipid metabolism, regulation, response to stress, structural and cell organisation and transport. Three genes involved with catabolic processes were only seen in clusters 1 and 5. Eight genes involved with cell replication were only seen in clusters 1, 2, 3 and 5. Three development genes were found in clusters 1, 2, 3. Six genes involved with the generation of precursor metabolites and energy were found in clusters 2, 4, 5 and 6. Four homeostasis genes were found in clusters 1, 3, 4, 5 and 6.

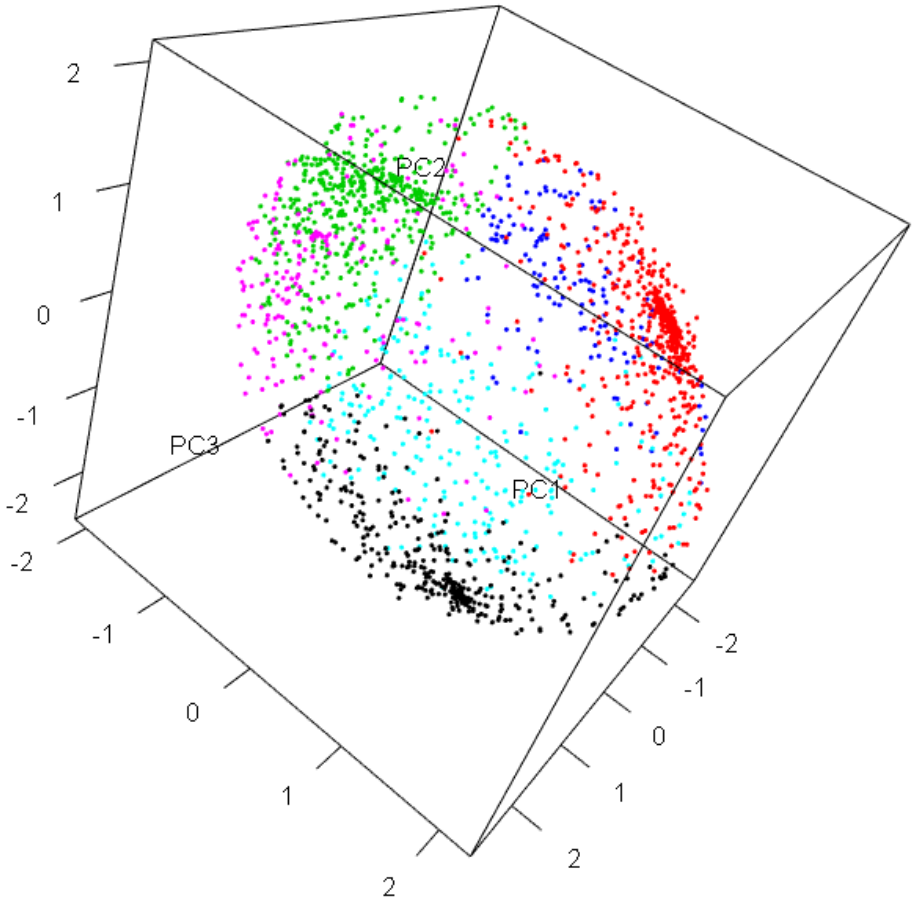


Figure 3.9 – The clustering of significantly differentially expressed genes by expression pattern analysed with principal component analysis. Significantly differentially expressed genes after exposure to SnTox3 grouped into six clusters with three principal components (PC) representing 83 % of the variation in expression patterns.

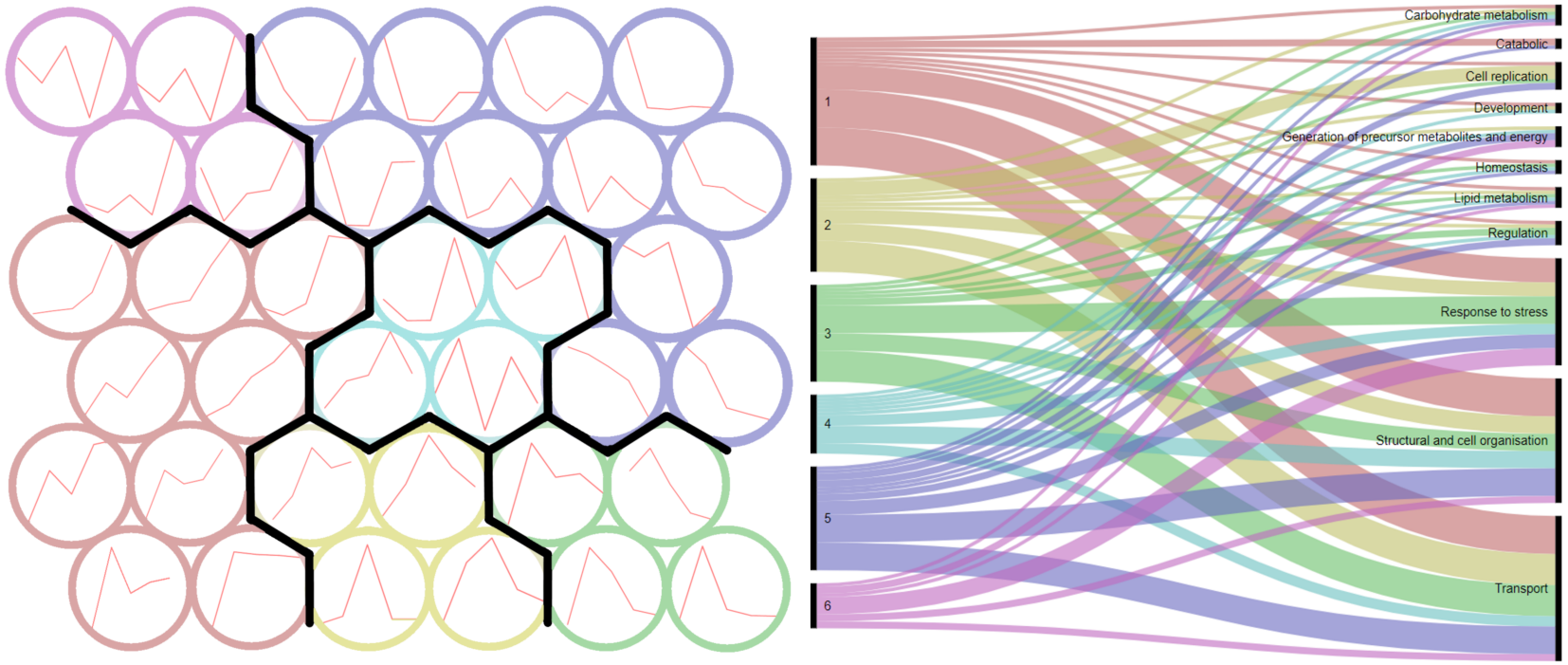


Figure 3.10 – The clustering of significantly differentially expressed genes by expression pattern analysed with GO Slim Terms. Significantly differently expressed genes (when comparing Cadenza SnTox3+ with all other samples; adjusted-P < 0.05) after exposure to SnTox3 grouped into 36 expression pattern types based on the five timepoints investigated (0, 4, 8, 12, 24 hours post infiltration) using self-organising maps, and then grouped into six clusters. These clusters were analysed for the gene ontology (GO) Slim Terms associated with the genes in each cluster.

3.4.7 Gene Structure and Predicted Protein Models of Candidate Genes Identified via RNA-seq Analysis

Intron/exon structure for g187 and g188 in the SnTox3 sensitive variety Cadenza were determined via comparison of Cadenza genomic sequence, the bioinformatically determined AUGUSTUS gene models, and the RNA-seq data generated in this chapter. Using Interpro to identify domains, protein models were subsequently developed. The CDS of g188 consists of five exons with the predicted protein possessing a signal sequence followed by a galacturonic acid binding (GUB)-WAK domain in the first exon. The fourth exon encodes an epidermal growth factor (EGF) domain and then a transmembrane domain is encoded by the end of the fourth exon and the start of the fifth exon. Finally, a S/TPK domain is encoded within the fifth exon (Figure 3.11). A second splice form of g188 was also present in the RNA-seq data which encodes a single exon with premature stop codon producing a protein of only 323 amino acids (aa) in length. The g187 CDS is six exons in length, with exons one to five encoding a S/TPK domain and exon six encoding a motile sperm domain. Additionally, a long 1,041 bp 5′ untranslated region (UTR) and a 685 bp 3′ UTR were identified (Figure 3.12). A second splice form was present for g187 consisting of four exons and a premature stop codon within the S/TPK domain. Interestingly, the genomic location of g188 is within the AM GWAS interval defined in Chapter 2, and g187 is residing close to the boundary (Figure 3.13).

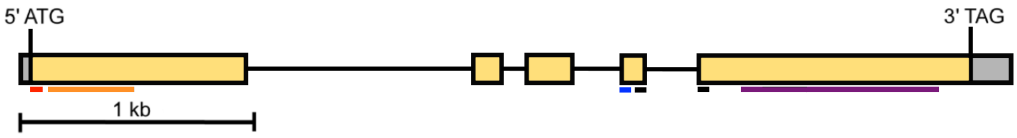


Figure 3.11 – Gene model of candidate g188 in cv. Cadenza. Grey indicates the untranslated region. The regions of the CDS predicted to encode protein domains are indicated as: red (signal sequence), orange (wall associated kinase domain), blue (epidermal growth factor domain), black (transmembrane domain), purple (serine/threonine protein kinase). The start (ATG) and stop (TAG) codons are indicated.

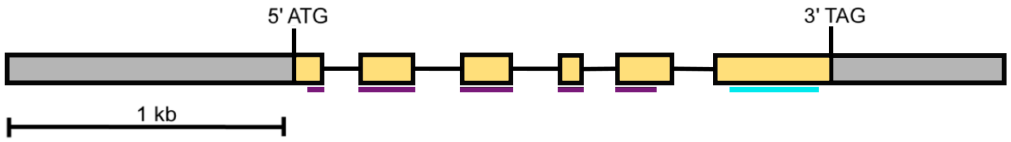


Figure 3.12 – Gene model of candidate g187 in cv. Cadenza. Grey indicates untranslated region. The regions of the CDS predicted to encode protein domains are indicated as: purple (serine/threonine protein kinase domain), light blue (motile sperm domain). The start (ATG) and stop (TAG) are indicated.

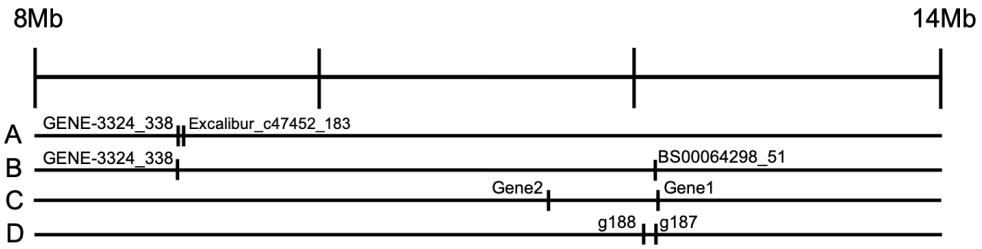


Figure 3.13 – An alignment comparing the relative locations of the candidate genes. The intervals displayed here are from the MAGIC SMA results (A) and the AM GWAS results (B) (Chapter 2), alongside the the candidate genes from Chapter 2 (C) and the candidate genes put forward in this chapter (D). These are aligned against the cv. Julius genome sequence. Showing Gene2 and g188 residing clearly within the AM interval and Gene1 and g187 sitting on the boundary.

3.4.8 TILLING

Infiltrating 1100 M_5 Cadenza TILLING lines with SnTox3 identified seven lines as having a change in phenotype from the wild-type sensitive phenotype (score = 4) to hyposensitive (score = 0). g188 was resequenced in all seven lines to identify mutations, as it is only 90–92 % similar to it's closest gene in Chinese Spring (TraesCS5B02G005755). For g187, TILLING mutations were identified using the TILLING data (available on plants.ensembl.org), as g187 is 97–100 % similar to it's most similar gene (TraesCS5B02G007900LC) in Chinese Spring IWGSC RefSeq (which is used in the TILLING database). Two mutations were identified in g188 in lines 287 and 779, of these one was synonymous and the other was intronic. Two mutations were identified in g187, both of which are non-synonymous. The other lines, had no identifiable mutations in g187 and g188.

Table 3.2 – Mutations within g188 and g187 in SnTox3 hyposensitive TILLING lines.

Candidate gene	TILLING line	SNP	SNP location	SNP type
g188	287	G → A	3611	synonymous Gly → Gly
	779	G → A	1282	intronic
g187	287	C → T	3511	non-synonymous Pro → Leu
	313	G → A	1975	non-synonymous Glu → Gly

Summary of SNPs present within candidate genes g187 and g188 in lines identified as SnTox3 insensitive (sensitivity score = 0) in the Cadenza TILLING population (Krasileva et al., 2017). Wild-type Cadenza control lines has a mean SnTox3 sensitivity score of 4.

3.4.9 Genomic Comparisons

The seven available scaffolded MAGIC founder assemblies (cvs. Alchemy, Brompton, Claire, Rialto, Robigus, Soissons, Xi19) were aligned to the Cadenza *Snn3-B1*

pseudomolecule via minimap2. Perfect matches (across all introns and exons) for both g188 and g187 were achieved with the Brompton to Cadenza alignment. However, g188 was found to be absent from the other six varieties' primary alignments to that region. g187 was represented in six alignments, but in five of these (Alchemy, Rialto, Robigus, Soissons, Xi19), there are noteable sequence differences between each MAGIC assembly and Cadenza with 16 SNPs, a 15 bp insertion and the first exon missing entirely. Both g188 and g187 were absent in Claire (Table 3.3).

Table 3.3 – MAGIC founder assembly alignments to the Cadenza *Snn3-B1* pseudomolecule

Founder	Mean SnTox3 Sensitivity	g188	g187
Alchemy	0	-	≈
Brompton	2	+	+
Claire	0	-	-
Rialto	4	-	≈
Robigus	0	-	≈
Soissons	4	-	≈
Xi19	4	-	≈

Summary of how the MAGIC assemblies align to the g187 and g188 candidate gene regions. + = present, ≈ = present with clear differences, - = absent.

To investigate whether g188 and g187 might be located on a MAGIC founder scaffold elsewhere on the chromosome, the raw reads from each MAGIC founder assembly were mapped to the Cadenza genome. This showed that again there was a perfect match between Brompton and Cadenza for both g188 and g187. But there were also MAGIC founder reads aligning to Cadenza g188 from Rialto, Soissons and Xi19, even though these were primarily in exonic regions only and had over 100 SNP differences between each MAGIC founder read set and Cadenza. Similarly it was also observed that g187 in Alchemy, Rialto, Soissons and Xi19 had a missing first exon and 16 SNPs. No reads from Claire and Robigus aligned to g188 and g187, additionally no reads from Alchemy aligned to Cadenza g188 (Table 3.4).

To further investigate the possible alleles of g187 and g188 present in wheat, the currently available wheat genome sequence assemblies were investigated. Accordingly, the presence of g188 (from Cadenza) and g187 (from Cadenza) as well the alternative versions identified in Rialto, Soissons and Xi19 (also Alchemy for g187) were investigated in an additional eleven wheat genomes (including the tetraploid Zavitan), and their SnTox3 sensitivity determined (Table 3.5). This showed that in SnTox3 insensitive lines (score < 2) both versions of g188 were absent. Intermediately SnTox3 sensitive lines ($2 \leq$ score < 4) have g188 version 1 (Cadenza-like) and highly sensitive lines (score $= 4$) have g188 version 2 (Rialto-like; Table 3.5). g188 shows clear clustering between the level of sensitivity and the gene

Table 3.4 – MAGIC founder raw read alignments to the Cadenza pseudomolecule

Founder	SnTox3	g188	g187
Alchemy	0	-	≈
Brompton	2	+	+
Claire	0	-	-
Rialto	4	≈	≈
Robigus	0	-	-
Soissons	4	≈	≈
Xi19	4	≈	≈

Summary of how the MAGIC assemblies align to the candidate gene regions. + = present, ≈ = present with clear differences, - = absent. g187 = serine/threonine protein kinase domain / motile sperm domain gene. g188 = wall associated kinase / epidermal growth factor / transmembrane / serine/threonine protein kinase gene.

version present, as well as abundance of gene entirely. g187, which also has two versions, Cadenza-like and Rialto-like, showed some degree of association with SnTox3 sensitivity. Varieties in which g187 is absent (Claire and Robigus) are insensitive to SnTox3. Those in which Cadenza-like g187 (v1) is present (Brompton, Paragon and Landmark) have intermediate sensitivity to SnTox3. But those that have the Rialto-like g187 (v2) fall into both the insensitive group and the highly sensitive group. Interesting, in those varieties that possess Cadenza-like g188 and g187, these two genes are located 1.2 kb apart, whereas those with rialto-like g188 and g187 are 170 kb apart. It is worth noting here that no other copies of g187 and g188 were found in these 170 kb regions. Due to the association of the Cadenza-like g187 and g188 with the intermediately SnTox3 sensitive varieties, the SnTox3 sensitivity was re-examinated in material, used in the RNAseq study, from the John Innes Centre (as this was used to sequence Cadenza) and was found to be intermediately sensitive (SnTox3 sensitivity score = 2) (Table 3.6; Figure 3.14). It is seen here that different Cadenza accessions exhibit different SnTox3 sensitivity scores; this has been reported previously with regard to PtrToxB sensitivity (Corsi et al., 2020).

3.5 Discussion

Very little is known about the SnTox3 response pathway and to date the identity of the major SnTox3 response gene, *Snn3-B1*, is unknown. However the genomic and RNA-seq datasets analysed in this chapter provide further understanding of the

Table 3.5 – Candidate genes present in each line

Variety	SnTox3 ± *s.e.*	g188v1	g188v2	g187v1	g187v2	distance between g188 & g187[†]
Alchemy	0 ± 0	-	-	-	+	
Chinese Spring	0 ± 0	-	-	-	+	
Claire	0 ± 0	-	-	-	-	
Jagger	0 ± 0	-	-	91 %	92 %	
Lancer	0 ± 0	-	-	-	+	
Robigus	0 ± 0	-	-	-	-	
SY-Mattis	0 ± 0	-	-	91 %	92 %	
Zavitan	0.17 ± 0.14	-	-	-	+ (99.6%)	
Arina	0.25 ± 0.14	-	-	-	+	
Brompton	2 ± 0.20	+	-	+	-	1213 bp
Paragon	2.3 ± 0.33	+	-	+	-	1213 bp
Landmark	2.5 ± 0.29	+	-	+	-	1213 bp
Julius	4 ± 0	-	+	-	+	170300 bp
Mace	4 ± 0	-	+	-	+	169451 bp
Norin61	4 ± 0	-	+	-	+	173320 bp
Rialto	4 ± 0	-	+	-	+	
Soissons	4 ± 0	-	+	-	+	
Xi19	4 ± 0	-	+	-	+	

Summary of which version of g188 and g187 each line has. P = present, - = absent. v1 = Cadenza-like, v2 = Rialto-like. g187 = serine/threonine protein kinase domain / motile sperm domain gene. g188 = wall associated kinase / epidermal growth factor / transmembrane / serine/threonine protein kinase gene.

[†] Distance is included where the genome sequence is pseudomoleculed or genes present on same scaffold.

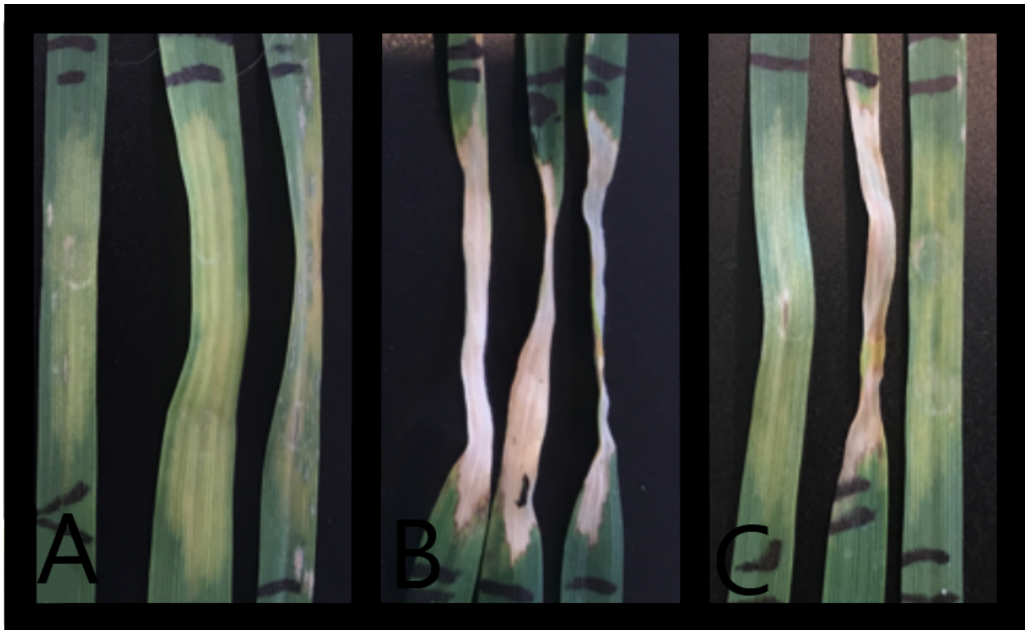


Figure 3.14 – SnTox3 infiltrations of cv. Cadenza from different sources. Cadenza, sourced from JIC, displaying an intermediate response to SnTox3 (A); Cadenza, from the AM panel (Chapter 2), exhibiting high sensitivity to SnTox3 (B) and Cadenza, the wildtype from the TILLING population, showing a mixture of high and intermediate sensitivity (C).

underlying pathway.

Table 3.6 – Cadenza accessions and their sources

Accession	Study	Chapter	SnTox3 sensitivity†	Source
Cadenza (AM panel)	GWAS	2	4 ± 0	NIAB DUS
Cadenza TILLING WT	TILLING	3	2.83 ± 0.60	GRU
Cadenza	RNA-seq	3	2 ± 0.29	GRU

Summary of how each Cadenza accession was used, it's corresponding SnTox3 sensitivity and it's source. GRU = Germplasm Resource Unit (John Innes Centre).
† SnTox3 sensitivity shown as mean ± *s.e.*

3.5.1 Minor SnTox3 Response QTL

Nine minor QTL for SnTox3 response were found in the Avalon × Cadenza and NIAB Elite MAGIC populations (Downie et al., 2018; Chapter 2). Here, RNA-seq found genes to be expressed within the genomic regions of three of these nine loci. However only two (*QTox3.niab-3B* and *QTox3.niab-7B*) show expression in the sensitive line, Cadenza. *QTox3.niab-3B* has one up-regulated DEG, TraesCS3B02G441300 (P < 0.0001; log2fold = 51.66), which encodes an amino acid permease. These are generally involved in amino acid transport systems, but can also be regulated by environmental stress (Kishor et al., 2005). Within the *QTox3.niab-7B* interval seven DEGs were identified, which are involved in a wide range of cellular processes. Four were significantly up-regulated and three were significantly down-regulated. The down-regulated DEG TraesCS7B02G032300 (P = 0.0037, log2fold change = -2.00), which is involved in protein detoxification, is also described as a multi antimicrobial extrusion (MATE) protein — a member of the MATE transporter family, which are reported to act as triggers for disease resistance pathways (Nawrath et al., 2002). The down-regulation of this gene may indicate that part of the plant's innate immunity is being suppressed in response to SnTox3 infiltration. TraesCS7B02G038100 (P = 0.15, log2fold change = 1.60) is also found in this locus, a pheophorbide a oxygenase, a class of gene known to have links with accelerated cell death (Pružinská et al., 2003). This may have a role in the programmed cell death of SnTox3 induced necrosis. TraesCS7B02G041100 (P = 0.0680, log2fold change = inf) is a DEG located in *QTox3.niab-7B* interval. This is a protein kinase, a class of gene well known to be involved in plant disease defense often linking with defense signalling cascades. Therefore the switching on of TraesCS7B02G041100 transcription identified here may indicate that SnTox3 has triggered a protein kinase cascade (Afzal et al., 2008). TraesCS7B02G045300 (P = 0.0590, log2fold change = -3.17), ferredoxin NADP reductase 2, another DEG located within the *QTox3.niab-7B* region, was down-regulated. These genes catalyse the reduction of NADP with reduced ferredoxin and are known to have a role in cellular defense against oxidative damage (Krapp et al., 1997). TraesCS7B02G050800 (P < 0.0001, log2 fold change = -1.91), a dirigent protein is involved in lignin biosynthesis, co-locates with *QTox3.niab-7B* and is

significantly down-regulated. Lignin biosynthesis deposits polymers in the plant cell wall, thus increasing the cell wall rigidity. This process has been known to be up-regulated in response to abiotic and biotic stresses (Vanholme et al., 2010). For example, the dirigent-like wheat protein, HfrDrd, is known to be involved in disease resistance in response to hessian fly larval attack (Subramanyam et al., 2013). However if significant down-regulation of this process occurs in response to SnTox3 infiltration, this could be an important factor on leading to necrosis, allowing for reduced resistance to cell breakdown. TraesCS7B02G061100 (P = 0.0002, log2fold change = 3.32), a Really Interesting New Gene (RING) finger protein, is up-regulated significantly, within the *QTox3.niab-7B* interval. RING finger proteins are known to mediate ubiquitin ligase activity (Joazeiro and Weissman, 2000). There are examples of pathogen induced RING finger proteins, such as CaRFP1, with roles in disease susceptibility and osmostic stress tolerance in *Capsicum annuum* (Hong et al., 2007). TraesCS7B02G065700 (P < 0.0001, log2fold change = 7.00) is a zinc finger protein and an up-regulated DEG in the *QTox3.niab-7B* region. Other zinc finger proteins, such as GhZFP1, have been seen to enhance fungal disease resistance and salt stress tolerance in *Gossypium hirsutum* (Guo et al., 2009). Both TraesCS7B02G061100 and TraesCS7B02G065700 may be involved in upregulated plant defense in response to SnTox3 infiltration. While it remains unclear how these DEGs may link in the SnTox3 response pathway in the susceptible variety Cadenza, the expression is statistically significantly increased/decreased in SnTox3 infiltrated Cadenza individuals. This indicates they may play a role in SnTox3 response, especially as there are examples where these classes of proteins have been involved or linked with disease resistance or plant defense. The potential role of the DEGs located within the physical intervals of the minor SnTox3 QTL discussed here is currently unknown, and further research is needed to clarify this. For the minor SnTox3 response QTL for which no DEGs were identified, there are many possible reasons for this: that they are not present in the Avalon or Cadenza genetic background; differential gene expression is not relevant to differences in allele function; the genes underlying these QTL do not have a SnTox3 specific role until after the first 24 hours; or that as the Chinese Spring genome was used to investigate the wider genome (beyond the *Snn3-B1* locus) these underlying genes are not annotated or are not present in this genome assembly.

3.5.2 Defense Related Pathways

Studies in *A. thaliana* have shown that many activated gene pathways and features are common to both PTI and ETI, such as hypersensitive response, ROS accumulation, *MPK3* and *MPK6* expression, salicylic acid, jasmonic acid and ethylene processes. However, there are a few key differences in these downstream responses, namely in

PTI ROS related genes, *MPK3* and *MPK6* are transiently activated, but in ETI their activation is far more prolonged (Tsuda and Katagiri, 2010). In this chapter, two candidate genes for *Snn3-B1* were identified: a WAK (g188) and a protein kinase (g187), located adjacent to each other in the Cadenza genome. WAKs have been reported as the product of resistance genes in wheat previously, for example *Stb6* which promotes resistance to *Zymoseptoria tritici*, also in a gene-for-gene manner (Saintenac et al., 2018). However, WAKs are more typical of a PAMP receptor, responsible for PTI initiation, and similar to Snn1 recognition of SnTox1. Whereas the protein kinase gene (g187), if it were to interact directly with SnTox3, would mean that the effector had been internalised and therefore it could trigger ETI, similar to Tsn1, responsible for recognition of SnToxA (Shi et al., 2016b). However, g188 also encodes a protein kinase and therefore could possibly act in either the PTI or ETI pathway. A key difference between recognition of SnToxA and SnTox1 is whether or not the effector is internalised or whether is it recognised in the apoplast by a cell wall associated receptor instead. Looking at this data in light of the nature of candidate genes g187 and g188, we can look for evidence of whether PTI or ETI is being hijacked in the host SnTox3 response pathways. Firstly, we can see here that ROS signalling and regulation is up-regulated throughout the 24 hour period after infiltration (Figure 3.7). Secondly, two of the *MPK3* homoeologues (from the A and B wheat genomes) are up-regulated immediately after infiltration, but not at later timepoints, even though the third *MPK3* homoeologue is up-regulated at 8 hpi. A proteomic study of SnTox3 infiltration in bread wheat found that MPK3 was also significantly induced at 24 and 48 hpi (Winterberg et al., 2014). Consistent up-regulation of ROS responses would suggest that ETI is activated. However the early up-regulation of MPK3 could be indicative of PTI. Although not seen in this study, the literature shows that MPK3 expression may also be prolonged, which is again indicative of ETI activation, making a stronger case for the ETI hijacking. At this stage, neither candidate gene can be excluded based on gene expression alone without further evidence.

However, the PR-1 proteins, that are significantly up-regulated in Cadenza after SnTox3 infiltration, are reported to display antifungal activity and play a role in host defense signalling and cell death and so have here been included with the antifungal proteins for analysis (Figure 3.7; Loon et al., 2006). Also it has been previously shown that PR-1 proteins interact directly with SnTox3 and SnToxA (Breen et al., 2016). In fact, it has been speculated that SnTox3, which is known to interact with PR-1, might mediate the release of CAPE1 (a defence signalling peptide from the C-terminus of PR-1 proteins) and in turn activate downstream defense signalling that involves or is hijacked by Snn3 (Breen et al., 2016).

Twelve of the 32 genes that were 'switched on' (expressed in Cadenza after

SnTox3 infiltration only) have known links with disease resistance, according to the KnetMiner database (Hassani-Pak et al., 2016). TraesCS4D02G202300 ($P = 0.0448$), which is transcribed after SnTox3 treatment in Cadenza, is a receptor-like kinase, linked with positive regulation of abscisic acid biosynthesis and therefore disease resistance. TraesCS4D02G221200 ($P < 0.0001$) is annotated as Topless Related Protein 1 (TPR1), which in other species is involved in disease resistance, gillerellin catabolic processes, regulation of starch biosynthesis, auxin biosynthesis. TPR1 is a known transcriptional corepressor of *Defense no Death 1* (*DND1*) and *Defense no Death 2* (*DND2*) in *A. thaliana*, two negative regulators of immunity that are repressed during pathogen infection. It is therefore possible that in wheat TPR1 activates R protein-mediated immune responses through repression of negative regulators in response to SnTox3 (Zhu et al., 2010).

The annotation of three of the twelve identified genes indicate links with plant response to stripe rusts: TraesCS3B02G366300 (protein kinase; $P = 0.0022$) and TraesCS3B02G366500 (protein kinase; $P = 0.0035$) with barley stripe rust resistance and TraesCS3B02G385500 ($P = 0.0020$) a V-type proton ATPase subunit a protein, linked with wheat stripe rust disease resistance. Two of the genes were identified as having possible links with necrosis: TraesCS3B02G049100 ($P = 0.0034$) and TraesCS5B02G544800 ($P = 0.0007$), both of which were predicted to encode receptor like kinases (Nekrasov et al., 2009; Larroque et al., 2013). The annotation of two genes, TraesCS2A02G577800 ($P < 0.0001$), TraesCS6B02G068400 ($P = 0.0031$), indicates association with disease-related metabolic processes, namely a glycosyltransferase, linked with disease resistance and lignin metabolism and a cytochrome p450 associated with CYP94D1, that catalyses two oxidation steps of the jasmonate pathway and disease resistance, respectively (Heitz et al., 2012). One of the 32 'switched on' genes, a terpene synthase linked to trehalose biosynthesis (TraesCS2A02G049500; $P < 0.0001$), has been found to be linked to *P. nodorum* sporulation (Lowe et al., 2009; Hassani-Pak et al., 2016).

Finally, TraesCS1B02G389700 and TraesCS1B02389900 were also identified as 'switched on' DEGs and were predicted to encode ethylene responsive transcription factors, a class of gene linked with camalexin biosynthesis and fungal defense, the latter also being linked with MAP kinase signalling and necrosis (Hassani-Pak et al., 2016).

3.5.3 Transcription Factors

Three genes encoding ethylene-responsive transcription factors (ERFs) were found to have an average normalised gene count of more than 26 fold compared to that of the second most expressed gene group at 0 hpi, the SPX domain-containing

genes, of the 'switched on' DEGs (Figure 3.8). These three ethylene-responsive transcription factors are adjacent gene neighbours. TraesCS1B02G389700 (ERF6), TraesCS1B02G389800 and TraesCS1B02G389900 and were expressed only in the SnTox3 sensitive cultivar Cadenza after exposure to SnTox3. Additionally, from 4 hpi onwards these three transcription factors are only expressed at a normalised gene count of one or fewer. They are likely to have a key role therefore in the early processes of SnTox3 response, triggered after Snn3 signalling, that leads ultimately to cell death, as, ERFs are linked with fungal defense and necrosis (Berrocal-Lobo and Molina, 2004). It has been found that down-regulation of *ERF6* leads to activation of the *Pathogen-inducible plant defensin 1.2* (*PDF1.2*) gene, which provides enhanced resistance to the necrotrophic fungus, *B. cinerea*, the causal agent of grey mould in *Vitis vinifera* (Amalraj et al., 2016). If *ERF6* down-regulation leads to enhanced resistance, it could be possible that up-regulation leads to enhanced susceptibility to another necrotroph, in a similar way to the observation that biotrophic and necrotrophic attack elicit upregulation of similar pathways, but with opposing effects (Kliebenstein and Rowe, 2008). It is therefore possible that these three *ERF* genes represent a key step in the hijacking of the wheat immune system by *P. nodorum*. In *A. thaliana*, mitogen-activated protein kinase 6 (AtMPK6), involved in the MAP kinase signal cascade, and is found to regulate gene expression of several key transcriptional factors, including ERF6, possibly prior to H_2O_2 accumulation (Liu and He, 2017). Interestingly, AtMPK6-ERF6 can also regulate jasmonic acid/ethylene-responsive genes in response to *B. cinerea* (Liu and He, 2017). The O_2 reacts to produce H_2O_2, but the latter can also be generated by peroxidases in the cell wall, e.g. in response to pathogen attack (Liu and He, 2017). This is perhaps a link with the peroxidases that are seen to be expressed following *ERF* transcription (Figure 3.8).

3.5.4 Clustering Analysis of DEGs

The clustering analysis shows that at the broader level all clusters contain very similar gene types, with a few exceptions containing or omitting one or two types (Figure 3.10). Interestingly, g188 belongs to cluster 3, which proportionally is one of the highest contributors to the category 'response to stress'. The closest 20 genes (in terms of euclidean distance within the PCA) to g188, which were analysed with more specific GO Terms, in this group include three with GO terms, 'response to toxin substance', three linked to 'regulation of plant-type hypersensitive response', two linked to 'defense response', 2 linked to 'response to wounding', one linked to 'MAPK cascade' and one linked with 'regulation of response to biotic stimulus'. If g188 is confirmed as *Snn3-B1*, these 20 nearest neighbours to g188 could be interesting targets for pathway analysis and further investigation.

3.5.5 *Snn3-B1* Candidate Genes

RNA-seq data was used to identify two candidate genes, designated g187 and g188. g187 has key RNA sequence differences between the SnTox3 sensitive and insensitive varieties and g188 is differentially expressed after exposure to SnTox3 in the sensitive variety Cadenza. While g188 may appear to be a promising candidate gene for *Snn3-B1*, as it appears at 4 hpi through to 24 hpi, it is unlikely to directly interact with SnTox3 given no expression was observed prior to infiltration. However, if it is assumed g188 is not expressed before infiltration, it may have a key role downstream as a wall-associated receptor kinase (Figure 3.3). In the genome assembly of the SnTox3 insensitive variety Chinese Spring, g188 is encoded by gene model TraesCS5B02G005755. This Chinese Spring gene has 90 % DNA identity, and is annotated as a nontranslating CDS, defined as a gene that appears to be protein coding but an invalid translation has been predicted. Indeed, this was confirmed by manual translation of the predicted CDS. For the Cadenza allele, g188, RNA sequencing showed it to result in a valid translation, predicted to encode a WAK domain, as also found in the SnTox1 sensitivity gene *Snn1*.

g187, however, is constituently expressed in both our SnTox3 sensitive and insensitive variety, but the sequence differs by four SNPs, an insertion and two large missing segments from both 5′ and 3′ ends of the RNA sequence. This is also a good candidate for *Snn3-B1* with a catalytic protein kinase domain, a domain that is common to both Tsn1 and Snn1. Based on the RNA-seq data generated here, the first 284 amino acids of Cadenza g187 are predicted not to be transcribed in Avalon g187, which contains a protein kinase domain, rendering g187 potentially non-functional or altering the function significantly. Additionally, the 1 bp insertion present in the Avalon is also predicted to result in a frame shift and premature stop codon based on direct comparison to the Cadenza allele. The intracellular protein kinase of Tsn1 is predicted to induce a MAP kinase cascade leading to cell death (Shi et al., 2016b). It is possible that g187 triggers ETS in a similar fashion and protein kinases are known to have a role in plant disease resistance, for example the barley stem rust R gene, *Rpg1*, contains a S/TPK (Brueggeman et al., 2002). Motile sperm domains, like the one identified here in g187 are typically involved in nematode sperm motility to provide a non-actin based way of movement that relays on protein-protein interactions (Smith and Ward, 1998). However the motile sperm domain has been previous reported in wheat, found as a domain encoded alongside NBS-LRR and integrated protein kinase domains in seven genes (Steuernagel et al., 2018). However aside with this report, the motile sperm domain is rarely reported with reference to plant genetics, although this link with NBS-LRRs suggests that they could have a role in plant immunity. Similarly, the EGF domain, found in g188,

is not a typical plant protein domain and is known for its role in blood coagulation, fibrinolysis, neural development and cell adhesion in mammals, however there is a small class of plant proteins known as the EGF-like receptor-like kinases, such as PRO25 in *Arabidopsis thaliana* (Campbell and Bork, 1993; Braun and Walker, 1996). Both g187 and g188 merit further investigation to understand their role in SnTox3 ETS.

Using the protein models determined in this chapter, the TILLING lines were examined for missense mutations and early stop codons that might explain the phenotypic change to hyposensitivity to SnTox3 (Table 3.2). The two missense mutations found in g187, do not code for a stop codon. However it is possible that they could fall in a critical amino acid residue. Alternatively, the SnTox3 TILLING mutants identified here via forward screening may be mutated at another gene in the SnTox3 sensitivity pathway. Determining whether TILLING mutations are allelic could be achieved by: (1) tests for dominance in F_1s generated by crossing to wild-type Cadenza (mutations would be assumed to be recessive in the absence of any other evidence), and then (2) by testing for complementation in combinations of F_1s between each mutant identified (i.e. if mutations were found to be recessive, combining two recessive mutations in the same gene would result in an SnTox3 insensitive F_1 individual. If recessive mutations were located in different genes, the resulting F_1 would have wild-type SnTox3 sensitivity). It is notable that in the 19 wheat genomes investigated, alleles at g188 perfectly associate with the three broad classes of SnTox3 sensitivity identified (insensitivity, intermediately sensitivity, high sensitivity), while alleles at the adjacent candidate gene g187 do not; Rialto-type g187 alleles were present in all of the highly sensitive varieties, as well as two insensitive varieties. However, the combinations of alleles between g187 and g188 meant that it is still possible that g187 could play a role in mediating SnTox3 sensitivity in varieties that carry functional sensitivity alleles at g188 (rather than a deletion of g188, as seen in all nine insensitive varieties for which genome assemblies are available). If this was so, identification of SnTox3 TILLING mutants in a highly SnTox3 sensitive Cadenza genetic background (the Cadenza germplasm from which the TILLING population is derived was found to sensitivity = 4) would be expected to be very rare, as homozygous deleterious mutations would be required at both g187 and g188 in the same TILLING line. If both genes do indeed play a role in SnTox3 sensitivity, then it may be more likely the insensitive Cadenza TILLING lines identified here via forward screening carry mutations at loci other than *Snn3-B1*, possibly explaining the lack of obvious deleterious mutations for g187 and g188 in the TILLING lines identified. Either way, the genetic basis for the SnTox3 insensitive Cadenza TILLING mutants identified requires further investigation.

The assembly to assembly alignments may not show sequence that is representative

of a particular region, as it maps each scaffold in its entirety. Therefore if a large part of the scaffold containing g188 maps well elsewhere, the entire scaffold will map elsewhere. Hence the raw read mapping provides a much clearer idea of whether g188 and g187 are present and if they differ in any way. We can see from the data that g188 alleles in Rialto, Soissons and Xi19 contain many SNP differences between them and g188 from Cadenza (Table 3.4). This holds true for g187 in Alchemy, Rialto, Soissons and Xi19. Therefore the Rialto version of each was extracted and compared with each available genome as the version two of these genes (Table 3.5). The clear clustering seen here between the versions of g188 (or absence of them) and phenotype give a very strong indication that g188 is *Snn3-B1*. It certainly would be a promising candidate for a perfect marker. g187 does show evidence of clustering between phenotype and genotype. However, it is plausible that the clustering of the intermediately sensitive lines with g187v1 is due to linkage, as in these cases the distance between g188 and g187 is only 1.2 kb, but when version 2 of g188 and g187 are present, the physical distance between the genes is approximately two orders of magnitude greater.

Interestingly, it is the Cadenza version of g187 and g188 that is found to be the intermediately sensitive version. The Cadenza lines phenotyped in Chapter 2 were sourced from NIAB, whereas the Cadenza lines used here in this chapter were sourced from the John Innes Centre's Germplasm Resource Unit and these lines had an average score of 2 (Table 3.6). It is likely that Cadenza used for the sequencing of the Cadenza genome was also sourced from the John Innes Centre. The differences in the SnTox3 sensitivity between the Cadenza accessions observed here is supported by similar reports of variation in Cadenza accessions elsewhere (Figure 3.14; Corsi et al., 2020). Indeed, when this is taken into account, analyses of g187 and g188 genomic sequences and SnTox3 sensitivity in the available wheat genome assemblies found: (1) at least three alleles are present for g188, and that absence is associated with insensitivity, the Cadenza allele is associated with intermediate sensitivity, and the Rialto allele is associated with high sensitivity. (2) Multiple alleles of g187 are present, and these show imperfect association with SnTox3 sensitivity. This indication that multiple alleles at the gene underlying *Snn3-B1* are present may explain why despite identification of the highly significantly associated SNPs being associated with SnTox3 sensitivity in the AM panel, those SNPs were far from being diagnostic. Indeed, in Chapter 2, combining the SNP call with presence/absence of Gene1 and Gene2 candidate genes improved genotypic prediction of SnTox3 sensitivity. Here, the investigation of g187 and g188 genomic sequences illustrates how moving beyond SNP markers to sequence/haplotype based analysis will likely result in rapid advances in the genetic dissection of agronomic traits in wheat.

In conclusion, this study identified two strong candidate genes for *Snn3-B1*, and

it is possible that allelic variation at one or both of these genes may play a role in the control of SnTox3 susceptibility. We have also seen that a number of genes have been differentially expressed in the minor SnTox3 sensitivity QTL *QTox3.niab-3B* and *QTox3.niab-7B*, some of which have known links to plant immunity. KASP markers for g187 and g188 have been designed for genotyping across the AM panel, but have yet to be validated. However given the strong bioinformatic evidence here, they may provide diagnostic markers for *Snn3-B1*. Furthermore, both candidates could be further investigated by producing knock-out mutants, individually or in combination, in a SnTox3 sensitive variety. In addition, we have seen the switching on of key metabolic processes / pathways linked with plant defense, linking with the hypothesis that *P. nodorum* uses SnTox3 to hijack plant immunity. In particular three *ERF* genes appear to play a central role in the wheat SnTox3 response pathway and these, alongside g187 and g188, are prime candidates for further investigation of this largely unknown pathway.

4

Wheat Genetic Loci Conferring Field Resistance to *P. nodorum* in European Germplasm

4.1 Abstract

Parastagonospora nodorum is a necrotrophic fungal pathogen of wheat (*Triticum aestivum* L.), causing Stagonospora Nodorum Blotch (SNB) on the leaves and ears. *P. nodorum* is known to produce proteinaceous necrotrophic effectors that result in the induction of host defense-related pathways that lead to cell death. Many field resistance quantitative trait loci (QTL) have been identified and some of these have been shown to co-locate with known effector sensitivity QTL. Here, using existing genotypic data, an improved genetic map for the Avalon × Cadenza bi-parental wheat mapping population, which is known to be polymorphic for SnToxA, SnTox1 and SnTox3 sensitivity, was developed. Utilising this denser and more accurate genetic map, a major QTL on chromosome 5A for both leaf and glume blotch, *QLb.niab-5A.2* and *QGb.niab-5A.2*, defined to an interval of only 12 Mbp, is identified. A further seven novel QTL for leaf or glume blotch were identified on chromosomes 1B (*QLb.niab-1B*), 2D (*QLb.niab-2D*), 3B (*QLb.niab-3B*), 5A (*QGb.niab-5A.1*), 5B (*QGb.niab-5B*) and 7B (*QGb.niab-7B*). Also in this study, field QTL co-locating with effector sensitivity QTL *Snn7* and possibly *Snn1* are identified.

4.2 Introduction

Parastagonospora nodorum is a necrotrophic fungal disease that, although commonly observed on wheat seedlings, has the greatest effect on the adult crop. It results in leaf blotch, which limits photosynthesis and ultimately crop growth, and/or glume blotch, which directly affects grain quality. It is important, therefore, to relate research of host genetic control of host sensitivity to *P. nodorum* effectors back to the genetic control of Stagonospora Nodorum Blotch (SNB) disease in the leaves and glumes in the adult crop.

Understanding the link between the constitutive genetic components controlling SNB resistance in the adult crop will inform wheat breeders and researchers alike, providing information on the overall phenotypic effects at the whole plant level, and allowing this to be related back to individual genetic interactions.

As described in Chapter 1, SNB results in chlorosis and necrosis of the leaf tissue, as well as discoloration and necrosis of the glumes, often in the form of lesions, referred to as leaf blotch and glume blotch, respectively (Solomon et al., 2006). Bread wheat has a global production of 728.2 million tonnes per annum, and is a key crop towards meeting the demands of a rapidly growing human population (Cui et al., 2014; Li et al., 2007; FAO, 2015). Improving the genetic resistance of wheat to the pathogens responsible for global yield losses is an important component of sustainable intensification approaches. SNB can cause wheat yield losses of up to 31 %, with SNB disease leading to significant economic loss in wheat crops across Scandinavia, the USA and Australia, as well as other wheat wheat producing regions such as southern Brazil (Bhathal et al., 2003; Eyal, 1999; Murray and J. P. Brennan, 2009; Hane et al., 2007). Much effort has already been made to identify sources of SNB resistance and effector sensitivity, with over 80 quantitative trait loci (QTL) identified to date (summarised in Tables 4.1 and 4.2, respectively). Although effector sensitivity studies provide useful insights, it is important that where possible these are related back to the adult SNB resistance. Therefore, in this chapter, where possible the data from Chapter 2 and other effector studies will be related back to the adult crop.

The Avalon × Cadenza population is the UK doubled haploid bread wheat bi-parental reference genetic mapping population, consisting of 206 progeny (Ma et al., 2015). The effector sensitivities of Avalon and Cadenza were determined in Chapter 2: Avalon displays insensitivity to SnToxA, SnTox1 and SnTox3 and Cadenza displays a necrotic response to all three effectors (R. C. Downie et al., 2018). Twelve wheat sensitivity loci to SnToxA, SnTox1 and SnTox3 have previously been identified (Table 4.2). However, the sensitivity of Avalon and Cadenza to SNB has not been determined to date. Therefore the Avalon × Cadenza population represents

a suitable resource for investigating adult plant response to *P. nodorum* infection and any potential correlation with effector sensitivity loci. However, while genetic maps for the Avalon × Cadenza population have previously been published, both available maps are limited in terms of marker density and chromosome fragmentation (Ma et al., 2015; Wilkinson et al., 2016).

4.2.1 Aims

While many QTL have been identified for seedling response to single effectors, this study uses the Avalon × Cadenza reference UK wheat bi-parental mapping population to identify QTL responsible for sensitivity to SNB in adult plants in a field environment and to potentially confirm and refine previously discovered SNB QTL, and to cross reference results with host sensitivity QTL for SnToxA, SnTox1 and SnTox3. Also a new Avalon × Cadenza map will be produced using the 90K array and a high quality eight founder genetic map as a back bone to attempt to maximise the single nucleotide polymorphism (SNP) marker density for this bi-parental mapping population.

Table 4.1 – *P. nodorum* Field QTL.

QTL	Type	Population	Markers	Location (Chr, Mbp)	PVE (%)	Reference
QSng.eth-1BS	Glume	Forno x Oberkulmer	xglk301, xglk317	1B	7.4	Aguilar et al., 2005
QSng.eth-2B			xpsr961, xglk687	2B	3.8	
QSng.eth-3A1			xpsr304, xpsr598	3A	0.3	
QSng.eth-3A2			xglk118, xglk577	3A	6.4	
QSng.eth-4A1			CD16.2, xglk331	4A	1.4	
QSng.eth-4A2			xglk128, xgwm710b	4A	4.3	
QSng.eth-4DL2			xglk302, xpsr1101a	4D	1.3	
QSng.eth-5A2			xpsr1194, xpsr918b	5A	35.8	
QSng.eth-5B1			xpsr128, xglk614	5B, 404.19	6.3	
QSng.eth-7AL			xglk165, pwir232a	7A	3.8	
QSnl.eth-1BS2	Leaf	Forno x Oberkulmer	xgwm18, xglk483	1B, 222.58	7.3	Aguilar et al., 2005
QSnl.eth-2A2			xpsr131, xglk687	2A	1.8	
QSnl.eth-2B1			xglk407, xpsr924	2B	8.1	
QSnl.eth-2B2			xpsr644b, xpsr956a	2B	4.7	
QSnl.eth-2D			xpsr932, xpsr331a	2D	20.0	
QSnl.eth-3B			xglk538, xpsr902	3B	9.9	
QSnl.eth-4B			xglk348, xpsr921	4B	17.0	
QSnl.eth-5B2			xglk163, xpsr426	5B	1.6	
QSnl.eth-7B1			xpsr952, xgwm46	7B, 158.082	0.9	
QSnl.eth-7B3			xmwg710a, xglk576	7B	12.8	
QSnl.ihar-2B	Leaf	Liwilla x Begra	xgwm120, xu36894	2B	16.0	Czembor et al., 2003
QSnl.ihar-5B			xbarc32, xgwm499	5B	30.0	
QSnl.ihar-5D			xgwm205, xgwm212	5D, 34–472	37.0	
QSnl.daw-1B	Leaf	05Y001	wPt8949, xgwm264	1B, 150.32–328.936	16	Francki et al., 2011

Continued on next page

Table 4.1 – continued from previous page

QTL	Type	Population	Markers	Location (Chr, Mbp)	PVE (%)	Reference
QSnl.daw-4B			xbarc0168, wPt0391	4B	8.2	
QSnl.daw-5B			wPt4628, wPt1733	5B	11.0	
QSnl.daw-2A		P9819RB1	wPt2448, wPt7056	2A	16.5	
QSnb.fcu-1BS	Leaf	BR34 x Grandin	xfcp267, xbarc240	1B, 389.99	12.0	Friesen et al., 2009
QSnb.fcu-2DS			xgwm614, xcfd53	2D, 23.024	12.0	
QSnb.fcu-4BS			xwmc47, xfcp301	4B, 644.866	7.0	
QSnb.fcu-5AL			xbarc151, xfcp13	5A, 558.34–625	14.3	
QSnb.fcu-5BL			xbarc1116, xbarc43	5B, 288.398	15.8	
SNG04	Glume	Synthetic Hexaploid Wheat	wPt3921, wPt8079		6.4	Jighly et al., 2016
SNG05			wPt7412		6.7	
SNL01	Leaf	Synthetic Hexaploid Wheat	wPt5069		7.6	Jighly et al., 2016
1B	Leaf	SHA3/CBRD x Naxos	xwmc619	1B, 65.282		Lu and Lillemo, 2014
3AS			xgwm2	3A, 60.201		
3B			wPt4127	3B		
5BS			wPt5346	5B		
5BL			xfcp1	5B, 546.147		
7A			xwmc603	7A, 488.729		
7B			wPt-0963	7B		
QSnb.cur-1BS	Leaf	Calingiri x Wyalkatchem	*Snn1*	1B, 2.354	0.19	Phan et al., 2016
QSnb.cur-6BS			wPt3168, xbarc146a	6B, 227.65	10.0	
1A	Leaf	SHA3/CBRD x Naxos	wsnp_Ex_c25734_34995416	1A, 472.168	4.4	Ruud et al., 2017
1B.1RS			SCM9	1B, 91.446	7.0	
3AS.1			xgwm2, IAAV6676	3A, 60.201–61.348	7.2	
3AS.2			Ku_c41007_116, Excalibur_c52446_519	3A, 537.508	6.1	
3BS.1			BS00030534_51	3B, 466.615	5.7	

Continued on next page

Table 4.1 – continued from previous page

QTL	Type	Population	Markers	Location (Chr, Mbp)	PVE (%)	Reference
3BS.2			wsnp_BE445348B_Ta_2_1	3B, 574.828	6.9	
3BL			wPt4933	3B	5.8	
5BS			BS00091518_51	5B, 6.648	11.9	
5B.2			wPt5914	5B	4.1	
7A			RAC875_c14195_1155	7A, 553.993	4.6	
7B			BobWhite_rep_c50229_413	7B, 328,802	5.6	
QSng.sfr-2AL	Glume	Arina x Forno	xcfd276a, xcfa2086	2A, 405.288	9.1	Schnurbusch et al., 2003
QSng.sfr-2BS			OA02, xpsr933b	2B	3.6	
QSng.sfr-2BL			xglk600, xpsr644b	2B	7.8	
QSng.sfr-3AS			xcfd79a, xgwm369	3A, 24.036–25.3656	7.7	
QSng.sfr-3BS			xgwm389, xcfd79c	3B, 0.806–18.84814	20.0	
QSng.sfr-3BL			xcfd2134b, xgwm131b	3B	9.0	
QSng.sfr-4BL			xgwm165, xglk335	4B, 509.036	13.2	
QSng.sfr-5AL			xglk317, xgwm639b	5A, 530	15.7	
QSng.sfr-5BL			xgwm371, xgwm639a	5B, 447.205–504.301	10.3	
QSng.sfr-5DL			xcfd81, xcfd266	5D, 44.567–375.854	7.3	
QSng.sfr-6BL			xfba81, xpsr924	6B	6.9	
QSng.sfr-6Dc			xcfd19, xgwm55b	6D, 62.078–283.599	5.0	
QSng.sfr-7BS			xcfa2174b, xgwm46	7B, 71.578–158.081	5.8	
QSng.daw-1B	Glume	WAWHT2074 x 6HRWSN125	xcfd59b, wPt3566	1B, 205.422	10.0	Shankar et al., 2008
QSng.daw-2A			wPt7026, xgwm71c	2A, 31.827	5.0	
QSng.daw-2B			xgwm148, wPt8072	2B, 100.848	8.0	
QSng.daw-2D			xcfd11, xgwm157	2D, 79.231–470.232	4.0	
QSng.daw-4B			Rht1, xgwm495	4B, 18.781–482.82	13.5	
QSng.daw-5A			xgwm617a, wPt0373	5A, 670.61	8.0	

Continued on next page

Table 4.1 – continued from previous page

QTL	Type	Population	Markers	Location (Chr, Mbp)	PVE (%)	Reference
QSng.daw-5B			wPt5896, xgwm604	5B, 385.994	3.0	
QSng.daw-7B			wPt4025, wPt6498	7B	5.0	
QSnl.daw-1B	Leaf	WAWHT2074 x 6HRWSN125	wPt3566, xcfa2129b	1B, 563.059	5.0	Shankar et al., 2008
QSnl.daw-2D			xcfd11, xgwm30	2D, 79.231–142.336	12.5	
QSnl.daw-4B			wPt3991, xbarc340	4B, 437.936	4.0	
QSnl.daw-5A			xgwm617a, wPt0373	5A, 670.61	21.0	
QSnl.daw-5B			wPt6136, wPt1420	5B	4.0	
3BS	Glume	Arina x Forno	swm01310bd	3B		Shatalina et al., 2014
3BS			wmm756	3B		
QSng-pur.2DL.1	Glume	P9819RB1	xgwm526.1, xcfd50.2	2D, 630.399-637.443	23.9	Uphaus et al., 2007
QSng-pur.2DL.2			xcfd50.3, wPt9848	2D, 637.443	9.4	

Summary of published *P. nodorum* field resistance QTL. PVE = Percentage variation explain.

† Physical location listed according to RefSeq BLAST hits, where marker sequence is available.

Table 4.2 – *P. nodorum* Effector QTL

QTL	Effector	Location (Chr, Mbp)	Markers†	Gene model	Reference
Tsn1	SnToxA	5B, 64.731		TraesCS5B02G059000	Faris et al., 2010
Snn1	SnTox1	1B, 2.354		TraesCS1B02G004100	Shi et al., 2016b
QSnn.niab-5A.1	SnTox1	5A, 702.461	BS00068108_51		Cockram et al., 2015
Snn3-B1	SnTox3	5B, 6.654, 6.647	Excalibur_c47452_183, GENE-3324_338		R. C. Downie et al., 2018
QTox3.niab-1B.1		1B, 349.715	Tdurum_contig76507_422		Chapter 2
QTox3.niab-2A.1		2A, 758.692, 758.554	BS00070979_51, Excalibur_c20478_641		R. C. Downie et al., 2018
QTox3.niab-2B.1		2B, 799.254	Kukri_c9898_1766		R. C. Downie et al., 2018
QTox3.niab-2B.2		2B, 771.205	Excalibur_c30571_95		Chapter 2
QTox3.niab-3B.1		3B, 67.942	wsnp_Ex_c11246_18191331		R. C. Downie et al., 2018
QTox3.niab-4D.1		4D, 32.347	BS00036421_51		R. C. Downie et al., 2018
QTox3.niab-6A.1		6A, 22.443, 22.524	BobWhite_c13839_135, IACX7801		R. C. Downie et al., 2018
QTox3.niab-6B.1		6B, 8.410	wsnp_Ku_c2119_4098330		R. C. Downie et al., 2018
QTox3.niab-7B.1		7B, 3.541	BS00022127_51		R. C. Downie et al., 2018
QSnb.cur-4BL		4B, 641.521	wmc413, wpt-730303		Phan et al., 2016
Snn3-D1		5D, 5.597	xcfd18		Zhang et al., 2011
Snn2	SnTox2	2D, 14.032	xgwm614, xbarc95		Friesen et al., 2007
Snn4	SnTox4	1A	xBG262267, xksum182.1		Abeysekara et al., 2011
Snn5	SnTox5	4B, 640.976	xcfd22, xwmc349		Friesen et al., 2012
Snn6	SnTox6	6A	xBE424987, xBE403326		Gao et al., 2015
Snn7	SnTox7	2D, 608.633-629.648	xcfd44, xgwm349		Shi et al., 2015

Summary of published wheat sensitivity QTL for *P. nodorum* effectors SnToxA, SnTox1 and SnTox3. Where possible the physical location of the gene or most significant genetic marker has been listed.

† Where the gene has been cloned, the gene model is listed (IWGSC RefSeq v1.1).

4.3 Materials and Methods

4.3.1 Wheat Germplasm and Genotyping

The Avalon × Cadenza bi-parental bread wheat population, along with a set of control lines, was used for an inoculated *P. nodorum* field trial. The Avalon × Cadenza population consisted of the two parents and 198 doubled haploid lines, and was sourced from the John Innes Centre, Norwich, UK. The SNB control lines Arina (moderately resistant), Brakar (highly susceptible), Jenga (resistant), Kuban (resistant), Naxos (moderately susceptible), SHA3/CBRD (highly resistant), Tarso (moderately resistant), TUI/RL4137 (highly susceptible), Xi19 (highly susceptible) and Zebra (moderately resistant) were also used (M. Lillemo 2019, personal communication). The Avalon × Cadenza population was previously genotyped using the Illumina iSelect 90,000 feature wheat SNP array (Wang et al., 2014), resulting in 10,941 polymorphic SNPs, available through CerealsDB (Wilkinson et al., 2016). The SNP data were ultilised to develop a genetic map using R/qtl (Broman et al., 2003) with a maximum recombination fraction of 0.35 and a minimum logarithm of the odds (LOD) score of 6. Where possible SNPs were anchored to specific chromosomes using data from the previously published 'NIAB multi-parent advanced generation inter-cross (MAGIC) Elite' genetic map (Gardner et al., 2016).

4.3.2 Field Trial and Phenotyping

The Avalon × Cadenza population was grown in hill plots (small plots sown 50 cm apart in rows, 40 cm between rows, 20 g seed per plot) at the Vollebekk Centre at NMBU (Norges milijø og biovitenskapelige universitet - the Norwegian University of Life Sciences), Ås, Norway (geographical coordinates: 59.658762 latitude, 10.754889 longitude), over two seasons. The first season's trial was sown on 10^{th} May 2016 and phenotyped in the summer of 2016. The second trial was sown on 14^{th} October 2016 and phenotyped in the summer of 2017. Each line was represented by two replicate plots plus four to six replicates of the control lines. Experimental design followed an α lattice, divided into 36 blocks of 12. The trial design was generated by researchers at NMBU using the statistical software CropStat (v7.2; IRRI, 2007). At late tillering / early stem elongation developmental stage, the plots were exposed to *P. nodorum* infected straw from the previous year's most susceptible varieties, grown at the same site, and water misted for 5 minutes every 30 minutes from 10am to 8pm from growth stage (GS) 30 (early stem elongation) onwards in order to promote infection (Zadoks et al., 1974). To protect against yellow rust the trials were treated with a selective fungicide Forbel 750 (Bayer Crop Science, a.i: Phenpropimorph 750 g $\cdot$ ha^{-1}) at 2–3 week intervals while the trials were mist-irrigated.

Heading date (days after sowing) and plant height (cm) were recorded at GS 59 (ear emergence complete). Primary leaf blotch severity (percentage coverage of the plot) was recorded at GS 65 (anthesis half-way complete), then again at one week intervals to monitor disease progression. Finally, at GS 73 (early milk development), glume blotch was also recorded (percentage coverage of the glumes).

4.3.3 Analysis

Phenotypic data was processed in order to account for the known confounding variables of height and heading date: disease score means were calculated from the line replicates for each phenotype. Residuals of these means were calculated in R using linear regression from height and heading data (R Core Team, 2013). The residuals were then used for QTL analysis along with the Avalon × Cadenza genetic map generated in this chapter, consisting of 1797 uniquely mapped SNP markers. QTL analysis was carried out using R/qtl using the EM algorithm. LOD thresholds for major and minor QTL were determined with 1000 permutations at the 1 % and 5 % significance level, respectively (Broman et al., 2003). The percentage variation explained (PVE) was calculated as

$$PVE = 1 - 10^{-(2/N*LOD)}$$

where N is the number of individuals included in the QTL analysis and LOD is the highest LOD score of the QTL.

Broad sense heritability (h^2) was calculated by estimating components of variation from REML whilst considering all features of the experiment design. Heritability was then estimated as

$$h^2 = \sigma_G^2/(\sigma_G^2 + \sigma_e^2)$$

where σ_G^2 is the genetic variation between line means and σ_e^2 is the error variance appropriate to those means. Calculations were carried out in GenStat (v19; VSN International, 2011).

4.4 Results

4.4.1 *P. nodorum* Susceptibility Phenotyping

The 198 lines of Avalon × Cadenza population were phenotyped for adult plant sensitivity to *P. nodorum* of the leaves and glumes at a field site in Norway during two consecutive summers (2016 and 2017). Additionally, plant height and heading date were recorded. The severity of host response was scored as a percentage leaf

blotch coverage of the individual plot as well as percentage glume blotch coverage, where 0 % indicates no visible disease presence and 100 % indicates full visible coverage of the target tissue (i.e. leaf or glume). Heritability for both primary leaf blotch and glume blotch was found to be similar (primary leaf blotch h^2= 25.1 %; glume blotch h^2= 22.0 %).

Leaf and glume blotch coverage were found to be correlated with height and/or heading date (Table 4.3). SNB phenotypes in the control lines were largely as expected with Jenga, Kuban and Sha3/CBRD displaying relative resistance (17.5 % [2017], 21.2 % [2017] and 15.0 % [2016], respectively). Arina, Tarso, and Zebra displayed moderate resistance (33.7 % [2017], 41.2 % [2017] and 31.0 % [2016], respectively) and TUI/RL4137 displayed high susceptibility (60 % [2016]). In 2016 and 2017, glume blotch scores were affected by height only. Leaf blotch from the second scoring date (subsequently termed here 'secondary scores') were affected significantly by both heading date and height in both years. In 2016, primary leaf blotch scores (i.e. from the first scoring date) were influenced by heading date, height and the interaction between these two confounding variables. However in 2017, primary leaf blotch scores were only significantly influenced by heading date. Tertiary leaf blotch scores were only recorded in 2016, and were influenced by both height and heading date. To account for these relationships, the residuals from the significant regression equations were used for further analysis. In 2016, height ranged from 45 cm to 73 cm with an average of 65.7 cm. Cadenza had an average height of 73.3 cm. Heading date ranged from 47 to 71 days with an average of 64.7 days. Cadenza had an average heading date of 63.3 days. Height and heading date for Avalon, in 2016, was not recorded as it failed to tiller, due to a spring sowing. In 2017, height ranged from 38 cm to 90 cm with an average of 62.3 cm. Avalon and Cadenza had an average height of 57.5 cm and 62.2 cm, respectively. Heading date ranged from 254 days to 270 days with an average of 260.9 days. Avalon and Cadenza had an average heading date of 260.5 and 259.8 days respectively. As the 2016 trial was spring sown Avalon (a winter wheat) failed to mature to heading.

4.4.2 Avalon × Cadenza Genetic Map Construction

While genetic maps using the Avalon × Cadenza population have previously been published (Ma et al., 2015) and the more recent map available via the CerealsDB online database (Wilkinson et al., 2016), both are limited in terms of chromosome fragmentation and marker density. Initial comparison with the 'NIAB Elite MAGIC' genetic map (Gardner et al., 2016) indicated it may be possible to improve the Avalon × Cadenza map. Accordingly, existing iSelect 90,000 feature SNP array data (Wang et al., 2014) for 198 Avalon × Cadenza doubled haploid lines, available via

Table 4.3 – *P. nodorum* coverage adjusted to remove confounding affects.

Measured Phenotype	Year	Significant Regression Equation
Primary leaf blotch score	2016	PLB ~ HD * Height
Secondary leaf blotch score	2016	SLB ~ HD + Height
Tertiary leaf blotch score	2016	TLB ~ HD + Height
Glume blotch	2016	GB ~ Height
Primary leaf blotch score	2017	PLB ~ HD
Secondary leaf blotch score	2017	SLB ~ HD + Height
Glume blotch	2017	GB ~ Height

% Summary of regression analysis to assess whether height or heading date (HD) are confounding variables for primary leaf blotch (PLB), secondary leaf blotch (SLB), tertiary leaf blotch (TLB) and glume blotch (GB) scores. Here the resulting significant regression equation is displayed for each phenotype.

CerealsDB, were used to create a new genetic map. Initially, 8 lines were removed due to high levels missing data (> 20 %) of missing data. Of the 10,941 polymorphic SNP markers for which there were available genotype data, 9322 had < 20 % missing data and were unskewed (defined here as the two homozygous genotypes sitting between 40-60 %). Accordingly, a total of 9322 markers were used for the construction of a genetic map, however, of these, 9266 markers were mapped to the 21 wheat nuclear chromosomes and the remaining 56 markers were unmappable (Figure 4.1). The markers per chromosome varied from 34 (chromosome 3D) to 1014 (1B). 39.4 %, 50.9 % and 9.8 % of markers were distributed across the A, B and D genomes, respectively. The map length summed 4838 cM, with single chromosomes ranging from 59.7 to 414.0 cM. The 9266 markers were located to 1797 unique sites across the genome, with 40.0 %, 47.4 % and 12.6 % of these unique sites belong to the A, B and D genomes, respectively. The map density was 0.37 unique sites per cM, and was 0.41, 0.44 and 0.20 for the A, B and D genomes, respectively (Table 4.6).

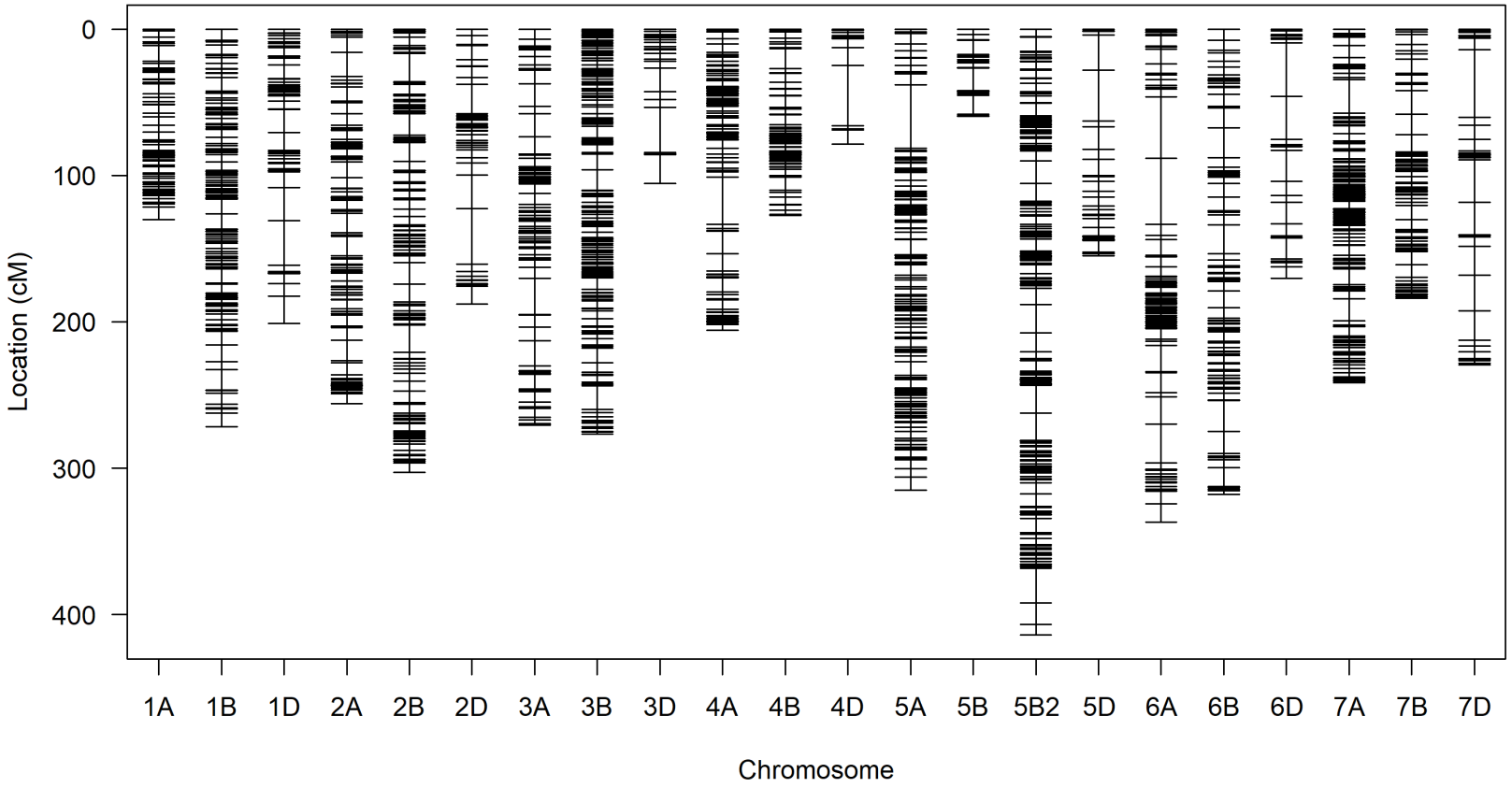


Figure 4.1 – The improved genetic map of the wheat Avalon × Cadenza doubled haploid population. Consisting of 9266 SNPs and 1797 unique marker positions, using 198 lines. Chromosome 5B is resolved into two separate, ordered, linkage groups: 5B and 5B2.

4.4.3 Genetic Analysis of *P. nodorum* Field Resistance

Using the new Avalon × Cadenza genetic map consisting of 9266 SNPs, QTL analysis was undertaken using the residuals of the collected SNB phenotypic data to correct the disease severity data against the confounding effects of height and heading date on leaf blotch disease severity. Using empirical 1 % and 5 % significance levels calculated using 1000 permutations to represent thresholds for 'major' and 'minor' QTL, respectively. Twelve QTL were identified in total (six major, six minor). Of these, four QTL co-located to very similar genomic regions, resulting in a total of 10 unique genomic loci (Table 4.4). These accounted for 7.5–18.3 % of the total variation (mean 13.4 %). Of the four different major QTL identified, one genomic locus, at which QTL for both leaf blotch (*QLb.niab-5A.2*) and glume blotch (*QGb.niab-5A.2*) were present, was found to be consistent between traits and years. This locus was located on the long arm of chromosome 5A with the peak location at 211.4 cM (SNP Ra_c3966_2205). The narrowest genetic interval of these co-locating chromosome 5A QTL, *QGb.niab-5A.2* for glume blotch in 2016, is located between 596.5 Mb (Ku_c12469_983, 210.6 cM) and 608.6 Mb (RAC875_c3964_752, 217.1 cM), and is predicted to contain 414 gene models (IWGSC RefSeq Annotation v1.1, representing 243 high confidence and 171 low confidence gene models). However the widest co-locating QTL at this genomic location, *QLb.niab-5A.2a* for primary leaf blotch, 2017, lies between 561.3 Mb (BS00067209_51, 177.4 cM) and 636.4 Mb (BS00104432_51, 227.27 cM) with 2127 gene models (947 high confidence and 1181 low confidence).

The other three major QTL were identified for glume blotch in 2016 and located on chromosome 5A at 239.4 cM (TA003720-0955), 248.8 cM (Tdurum_contig11521_102) and on chromosome 5B2 at 27.9 cM (Tdurum_contig30677_66), contributing between 13.9–15.5 % of the phenotypic variation (Table 4.4). The remaining six minor QTL are located on chromosomes 1B (8.6 cM, tertiary leaf blotch 2016; 33.0 cM, glume blotch 2016), 2D (168.7 cM, primary leaf blotch 2017), 3B (13.0 cM, secondary leaf blotch 2017), 5A (138.7 cM, glume blotch 2016) and 7B (107.9 cM, glume blotch 2016). The six minor QTL explain between 7.5–15.8 % of the phenotypic variation. No QTL were identified for primary and secondary leaf blotch in 2016 or glume blotch in 2017.

In addition to the twelve leaf/glume blotch QTL identified, seven and eight QTL were identified for heading date and height, respectively (Table 4.5). No overlap was observed between QTL for leaf or glume blotch and those for height or heading date.

Table 4.4 – QTL analysis of leaf and glume blotch in the Avalon × Cadenza bi-parental population grown in field trials in Norway and phenotyped during the summers of 2016 and 2017.

Phenotype	Year	QTL	Maj/Min†	Peak (Chr, cM, Mbp)	Marker	LOD	Interval (cM)	PVE (%)
Glume blotch	2016	*QGb.niab-1B*	Min	1B, 33.0, 50.184	IACX502	3.46	30.3–42.4	13.6
		QGb.niab-5A.1	Min	5A, 138.7, 533.071	BobWhite_c19155_246	3.43	136.2–154.0	13.5
		QGb.niab-5A.2	Maj	5A, 211.4, 596.465	Ra_c3966_2205	3.76	210.6–217.1	14.7
		QGb.niab-5A.3	Maj	5A, 239.4, 657.192	TA003720-0955	3.58	236.6–245.0	14.0
		QGb.niab-5A.4	Maj	5A, 248.8, 665.779	Tdurum_contig11521_102	3.55	247.8–249.8	13.9
		QGb.niab-5B	Maj	5B2, 27.9, 606.845	Tdurum_contig30677_66	3.99	22.31–33.56	15.5
		QGb.niab-7B	Min	7B, 107.9, 712.073	BS00110528_51	3.02	105.32–109.06	12.0
Primary leaf blotch	2017	*QLb.niab-2D*	Min	2D, 168.7, 637.653	Kukri_c498_2381	3.25	165.65–171.21	7.46
		QLb.niab-5A.2a	Maj	5A, 211.4, 596.465	Ra_c3966_2205	8.49	177.4–227.27	18.34
Secondary leaf blotch	2016	*QLb.niab-3B*	Min	3B, 13.0, 5.954	RAC875_rep_c106279_291	3.76	11.48–14.79	8.54
		QLb.niab-5A.2b	Maj	5A, 211.4, 596.465	Ra_c3966_2205	6.15	185.91–223.18	13.58
Tertiary leaf blotch	2016	*QLb.niab-1B*	Min	1B, 8.6, 1.779	BS00076192_51	3.95	0–17.31	15.77

Summary of major (Maj) and minor (Min) QTL found in the Avalon × Cadenza population using leaf blotch and glume blotch residuals. PVE = percentage variation explained. QTL replicated across both trials are highlighted in bold.

† 'Major' and 'minor' QTL represent loci above in the 1 % and 5 % empirical significance threshold determined using 1000 permutations.

Table 4.5 – Summary of QTL identified for plant height and heading date in the Avalon × Cadenza population grown in field trials during 2016 and 2017 in Norway. QTL replicated across both trials are highlighted in bold.

Phenotype	Year	QTL	Peak (Chr, cM)	Marker	LOD	Interval (cM)
Height	2016	*QHeight.niab-2A*	2A, 255.8	Excalibur_c20448_318	3.41	246.68–255.83
		QHeight.niab-2D.1	2D, 25.0	JD_c63957_1176	3.25	20.93–33.14
		QHeight.niab-4D	**4D, 4.3**	**BobWhite_c43880_73**	**8.02**	**0.00–5.44**
		QHeight.niab-4D.2	**4D, 25.0**	**tplb0038a13_900**	**11.3**	**12.30–66.12**
	2017	*QHeight.niab-2D.2*	2D, 33.1	GENE-0717_28	8.00	11.16–57.92
		QHeight.niab-3A	3A, 148.7	RAC875_c15003_377	3.47	145.91–150.00
		QHeight.niab-4D.3	**4D, 6.0**	**BS00031366_51**	**3.88**	**2.29–12.60**
		QHeight.niab-4D.4	**4D, 25.0**	**tplb0038a13_900**	**4.27**	**12.60–66.12**
Heading date	2017	*QHD.niab-1B*	1B, 191.0	RAC875_c64377_732	4.80	188.78–198.47
		QHD.niab-1B.2	1B, 248.6	Tdurum_contig8669_296	3.56	247.04–256.24
		QHD.niab-1D.1	1D, 86.2	BS00082503_51	3.96	84.98–88.56
		QHD.niab-1D.2	1D, 161.5	Kukri_c29687_369	7.85	108.22–201.08
		QHD.niab-2D.1	2D, 60.5	Kukri_rep_c114322_506	3.91	59.45–61.52
		QHD.niab-2D.2	2D, 77.5	wsnp_Ex_c25311_34578436	4.11	76.23–79.32
		QHD.niab-6A	6A, 199.4	BobWhite_c6950_269	3.81	198.31–202.43

Summary of QTL found in the Avalon × Cadenza population for height and heading date.

4.5 Discussion

In many regions of the world, SNB is a major disease of wheat. As resistance to SNB is quantitative in nature, and gains in wheat resistance relies on a system of marginal gains, breeding for improved resistance can be a challenging process. Currently identification of necrotrophic effectors allows host susceptibility to be dissected into some of its constitutive parts. However field studies of SNB do not always reflect this, as the co-location of effector sensitivity and SNB QTL are not always identified, although SnToxA is known to significantly correlated with increased SNB severity (Faris et al., 2011; Ruud et al., 2017). Therefore field resistance relies on multiple minor effect genes (Cockram et al., 2015). As resistance to SNB currently largely relies on a system of marginal gains, this chapter focuses in particular on a major QTL for field resistance identified using the improved Avalon × Cadenza genetic map developed here. Of the 82 previously identified field resistance QTL listed in Tables 4.1, only five have been linked to previously published effector susceptibility QTL, typically studied in seedlings: *QSnb.fcu-5BL* (Friesen et al., 2009) and *5BL* (Lu and Lillemo, 2014) with *Tsn1*; *QSnb.cur-1BS* (Phan et al., 2016) with *Snn1*; and *5BS* (Lu and Lillemo, 2014) and *5BS* (Ruud et al., 2017) with *Snn3-B1*. Either many more effectors remain undiscovered, their epistatic interactions are not sufficiently understood, or there are many more factors involved in building SNB resistant varieties. In fact the phenomenon of stage specific resistance should not be ruled out in SNB either, as this has been identified as key resistance genes to the rusts and powdery mildew in bread wheat are found to be effective at the adult stage but not at the seedling stage (Rinaldo et al., 2017).

4.5.1 The New Avalon × Cadenza Genetic Map

The new Avalon × Cadenza genetic map produced here provides a denser genetic resource, than the previously published map, by 37 % (Table 4.6; Ma et al., 2015). In fact, this enriched overall marker density is also 6 % denser than the CerealsDB Avalon × Cadenza genetic map (Wilkinson et al., 2016), possible due to using a high density eight founder genetic map as a backbone for the map construction. As the updated map generated here is comprised of SNP markers from the popular iSelect 90,000 feature SNP array, it allows for easy anchoring of the genetic map to the wheat physical map constructed in cv. Chinese Spring, as well as to other genetic maps and genotype data generated using the same array, whilst maintaining (and in indeed exceeding) the high standard set by CerealsDB. For example, the Rialto × Savannah population (available on CerealsDB), a wheat consensus map made from six bi-parental populations (Wang et al., 2014), SHA3/CBRD × Naxos (Ruud et al., 2017), the 'NIAB Elite MAGIC' population (Gardner et al., 2016) and

association mapping (AM) panels (R. C. Downie et al., 2018). This is in contrast with the Ma et al. (2015) Avalon × Cadenza map, developed with a combination of SNP, simple sequence repeat (SSR), diversity arrays technology (DArT) and perfect markers, some of which have limited public data on genomic location, meaning that comparison between genetic maps is not always straight-forward — although incorporation of these additional marker types does help cross-comparison with older genetic maps. The generation of high quality genetic maps greatly helps downstream genetic mapping activities. The refined Avalon × Cadenza map generated here provides a useful resource to help downstream genetic mapping activities conducted with this UK reference bi-parental mapping population. It is hoped that this map will be of use to other researchers using this population in the future.

Table 4.6 – Comparison of genetic maps available for the Avalon × Cadenza population.

	Ma et al., 2015	**CerealsDB**‡	**New map**
Number of lines	202	198	198
Number of linkage groups	28	23	22†
Number of markers	855	11029	9266
Number of unique markers	855	2057	1797
Total length (cM)	3188.28	5867.3	4838
Average linkage group length	113.9	255.1	219.9
Overall density	0.27	0.35	0.37
A genome			
Number of unique markers	276	820	718
Total length	1058.66	2221.30	1756.21
Density	0.26	0.37	0.40
B genome			
Number of unique markers	365	957	852
Total length	1180.29	2214.7	1954.16
Density	0.31	0.43	0.44
D genome			
Number of unique markers	211	280	227
Total length	949.33	1431.3	1127.55
Density	0.22	0.19	0.20

Comparison of summary data for publicly available genetic maps for the Avalon × Cadenza double haploid mapping populations with that of the new Avalon × Cadenza genetic map generated here in this chapter.

† chromosome 5B is represented by two ordered, orientated linkage groups: 5B and 5B2.

‡ map available from cerealsdb.uk.net

4.5.2 Analysis of the Co-locating Leaf and Glume Blotch *QLb.niab-5A.2* and *QGb.niab-5A.2*

One consistent major QTL for both leaf and glume blotch was found on the long arm of chromosome 5A in both 2016 and 2017 (Table 4.4). Three additional major QTL were found on chromosomes 5A and 5B for glume blotch in 2016. A SNB

resistance QTL was first reported on chromosome 5A in 1993 in a 5A substitution line (Nicholson et al., 1993). Later, a glume blotch resistance QTL, *Qsng.eth-5AL*, was identified on the long arm of chromosome 5A in the bi-parental mapping population Forno × Oberkulmer, accounting for 36 % of the variation (Aguilar et al., 2005). In 2009, another QTL was mapped in this region (*QSnb.fcu-5AL*), contributing to leaf blotch resistance in the Grandin × BR34 population (Friesen et al., 2009). Anchoring markers associated with *QSnb.fcu-5AL* (xfcp13 and xbarc142) to the wheat reference genome (IWGSC RefSeq v1.0; Appels et al., 2018) locates this region to 558–605 Mb on chromosome 5A, which coincides with the major *QLb.niab-5A.2*/*QGb.niab-5A.2* locus identified here (peak marker: Ra_c3966_2205; 596.5 Mb, and localised to a 12 Mb region in Avalon × Cadenza). Importantly, *VRN-A1*, a gene controlling the transition from vegetative to reproductive meristem development in response to vernalisation, is located on chromosome 5A at 587.42 Mb. Although *VRN-A1* lies within the *QLb.niab-5A.2* interval in 2017, and Avalon and Cadenza are known to carry contrasting alleles at this locus, *VRN-A1* is located outside the *QGb.niab-5A.2* interval (Ma et al., 2015). Furthermore, while it does not appear that *VRN-A1* represents the underlying gene, it should not be ruled out as a candidate.

Of the 82 previously reported field QTL listed in Table 4.1, there are only three co-locating glume-leaf blotch QTL pairs: *QSnl.eth-2B2* and *Qsng.sfr-2BL* (via xpsr644b); *Snl.daw-2D* and *QSng.daw-2D* (via xcfd11) and *QSnl.daw-5A* and *QSng.daw-5A* (via xgwm617a and wPt0373; Tables 4.1; Schnurbusch et al., 2003; Aguilar et al., 2005; Shankar et al., 2008). Therefore, it would appear that co-localisation of leaf and glume blotch QTL in the published record is relatively rare, hence the co-locating QTL *QLb.niab-5A.2* and *QGb.niab-5A.2* identified here may represent an important locus conferring resistance to both *P. nodorum* diseases in European germplasm.

Taking a broader look at crop fungal pathogens, the general dissociation between glume and leaf blotch QTL demonstrated in the SNB QTL recorded thus far is not a feature unique to SNB. In fact specific tissue infection has been well documented in the rice blast fungi, *Magnaporthe oryzae* and *Magnaporthe grisea* (Dong et al., 2015; Dufresne and Osbourn, 2001). In this case deletion of *MoMyb1* in *M. oryzae* gives raise to infection in rice roots, but not in the leaves, possibly due to changes in conidiogenesis, thus having an effect on the pathogen's basic life cycle (Dong et al., 2015). If rice blast here is taken as a case study for mechanisms that achieve tissue-specific infection, then perhaps in SNB the often dissociated leaf and glume QTL show the flip side of the coin. These QTL may be key in providing resistance at specific points in the infection process. Thus by identifying and maintaining co-locating leaf and glume blotch QTL, presumably with the same underlying genes, broader resistance to SNB's infectious stage may be achieved.

4.5.3 Analysis of the Nine Additional QTL Identified

The remaining nine SNB QTL identified within this study lie on chromosome arms 1BS, 2DL, 3BS, 5AL and 5BL. The PVE of these loci ranges from 7.5 to 15.8 %, although none of these QTL were replicated between years. QTL for *P. nodorum* resistance have previously been identified on all of these chromosome arms. However, being situated on the same chromosome is, of course, no guarantee that these represent the same loci. For example, in this study a glume blotch QTL, *QGb.niab-1B*, was identified on 1B at 32.95 cM with peak marker IACX502, and anchored to the wheat reference genome at 50.184 Mb. The gene underlying the major SnTox1 sensitivity locus *Snn1* has previously been cloned, and is location on chromosome 1B at 2.354 Mb (Shi et al., 2016b). Therefore, *Snn1* does not co-locate with *QGb.niab-1B*, identified here. However, four leaf and glume blotch QTL were found to co-locate with known SNB resistance QTL or major and minor effector sensitivity QTL: (1) Leaf blotch QTL *QLb.niab-2D* on chromosome 2D (peak marker: Kukri_c498_2381 at 168.7, 637.653 Mb) is located in a similar region to the SnTox7 sensitivity locus *Snn7* (Shi et al., 2015), based on the anchoring of the *Snn7* flanking markers Ra_c19051_1446 (638.147 Mb) and Excalibur_c42413_422 (638.376 Mb). Recently, Lin et al. (2020) used the 'NIAB Elite MAGIC' population to identify a seedling resistance QTL to two *P. nodorum* isolates that maps to this region (*QSnb.niab-2D.2* explaining up to 11 % of the variation, QTL peak between 636–638 Mb). Additionally, flanking markers for previously published glume blotch QTL (Francki et al., 2011; Uphaus et al., 2007) have also been noted as being located in the same region of chromosome 2D (Lin et al., 2020). Shi et al. (2015) suggested that the chromosome 2D glume blotch resistance QTL *QSng.pur-2DL.1* identified by Uphaus et al. (2007) was not *Snn7*, as neither of the parent lines of the population used were SnTox7 sensitive. Either way, it is clear that allelic variation at this region of chromosome 2D appears to be important for *P. nodorum* resistance in bread wheat. (2) leaf blotch QTL *QLb.niab-1B* at 8.6 cM on 1B (10–19 Mb) co-located with the previously identified leaf blotch QTL *QSnl.daw-1B* (Francki et al., 2011), anchored here to the wheat genome at 15 Mb via marker xgwm264. However, while these loci at first appear to be distinct from the *Snn1* locus (2.354 Mb), conferring sensitivity to SnTox1, the 1B linkage group does not represent any SNP before 10Mb so in fact this locus could co-locate with *Snn1* and *QSnb.cur-1BS*. (3) Glume blotch QTL *QGb.niab-5A.1* on chromosome 5A, anchored to the wheat genome between 510–548 Mb co-located with the previously identified leaf blotch QTL, *QSng.sfr-5AL* (Schnurbusch et al., 2003), based on the anchoring of marker xgwm639b to 530 Mb. (4) The location of the glume blotch QTL *QGb.niab-5B* (chromosome 5B at 606.845 Mbp) overlapped with the previously reported SNB QTL *5BL* (Lu and Lillemo, 2014), anchored at

546.147 Mbp using the physical map location of the peak marker, fcp1. As flanking markers for this QTL are unknown, limited conclusions can be made relating to whether these QTL represent the same locus. The remaining six QTL identified in this study, *QGb.niab-1B*, *QLb.niab-2D*, *QLb.niab-3B*, *QGb.niab-5A.3 QGb.niab-5A.4* and *QGb.niab-7B*, did not overlap with known SNB or effector sensitivity loci and appear to be novel QTL. However these QTL were only identified in a single year.

The environmentally sensitive QTL identified here, i.e. the QTL observed in a single year, are not an uncommon feature for phenotypes with a strong genotype × environment interaction, such as leaf blotch (Scott et al., 1982). Although the wheat population in this study remains the same from year to year, the *P. nodorum* population is known to be a highly genetically diverse population, due to its high levels of sexual reproduction (McDonald and Linde, 2002). High diversity levels have even been reported in isolates originating from the same SNB lesion (McDonald et al., 1994).

It is worth noting that of the overlap between QTL found in this study and previously reported QTL, only two were reported in seedling studies. This may have a few key underlying reasons. (1) The correlation between seedling and adult plant resistance is relatively low in the literature and therefore this is not unexpected (Ruud et al., 2018; Shankar et al., 2008; Uphaus et al., 2007). (2) Few seedling studies explore SNB resistance, but rather specific effector resistance. (3) When seedlings are exposed to SNB, this typically is carried out in an artificial glasshouse setting and therefore suggesting that the isolates in use may be quite different to those in the field. Particularly, like in this study, when the natural field population of *P. nodorum* is used (Downie et al. submitted).

Comparison with published SNB QTL indicates a number of novel QTL were found in this study. For those QTL validated across two seasons, further work is needed to refine the genetic intervals and further characterise the phenotypic effect of the QTL in isolation — for example, near isogenic lines (NILs) could be made for specific QTL, allowing detailed phenotyping to be undertaken, as well as fine-mapping of the Mendelised locus via crossing the NIL pair and identifying offspring with genetic recombinations within the QTL interval. For those QTL identified in just a season, validation in at least one additional season would be required in order for them to be prioritised for further investigation.

In conclusion, in this chapter while numerous QTL were identified for leaf and glume blotch, none of these co-located with sensitivity QTL for the known effectors SnToxA and SnTox3. Although, the chromosome 1B leaf blotch QTL identified here is in very close proximity to *Snn1*, however further markers and/or QTL mapping of SnTox1 sensitivity would be needed here to confirm. This indicates that genetics behind the *P. nodorum* field resistance is broader than the susceptibility

loci previously identified in Chapter 2 and in the wider literature for known effectors. Lastly, a co-locating QTL for both glume and leaf blotch has been identified here, suggesting that these *P. nodorum* mediated diseases have some genetic resistance components in common.

5

Discussion

5.1 Potential Further Work

The first priority would be to validate the identity of the gene(s) underlying *Snn3-B1* via reverse genetics approaches. The genomic and transcriptomic evidence presented here points to g188, a wall associated kinase (WAK), and this should be followed up with reverse genetic validation approaches such as clustered regularly interspaced short palindromic repeats (CRISPR) gene editing in a SnTox3 sensitive genetic background (Table 3.5). Knock-outs for g188, g187 and g188/g187 in combination in Cadenza and Rialto followed by SnTox3 infiltration of the T_0 and T_1 plants would clearly show if either g188 or g187 alone is response for SnTox3 sensitivity. The ability to also knock-out both candidate genes in a single genetic background would be important in case both are required within the SnTox3 response pathway that ultimately triggers hypersensitive response. This should be carried out, if possible, in Cadenza as most primary data here have been examined within this cultivar and protein models have been developed for Cadenza also. However, it should be noted that evidence of two contrasting accessions of Cadenza may exist in public genebanks, which have either medium (e.g. accessions from JIC SeedStor, assumed to be the source of the Cadenza genome assembly, and the Cadenza accession used for RNA-seq) or high (e.g. accessions used in the AM panel and A×C population) or indeed mixed (e.g. the Cadenza TILLING wild type) SnTox3 sensitivity. Evidence for unfixed alleles in Cadenza during the period it was used for breeding has recently

been noted via analysis of the transmission of sensitivity alleles to the *Pyrenophora tritici-repentis* effector ToxB across a large wheat pedigree (Corsi et al., 2020). Therefore, it would be important to confirm the candidate gene alleles present in the germplasm stocks used in any reverse genetic functional validation studies.

If possible and if financial constraints allow, these candidate genes should also be functionally validated using a highly sensitive variety, such as the MAGIC founder Rialto or the Australian variety Mace, as the genome sequences for the MAGIC founders will soon be available as part of a wheat pan-genomics project (gtr.ukri.org/projects?ref=BB%2FP010741%2F1) and that of Mace and other varieties will soon be made available via the 10+ Wheat Genomes Project. Functional investigation of the two candidate genes to ascertain whether knocking out g187 and g188, alone and in combination, in both an intermediately sensitivity and a highly sensitivity variety could provide additional insight into genetic control of the SnTox3 sensitivity by the *Snn3-B1* genetic locus.

Following this, further functional analysis on the confirmed *Snn3-B1* gene could be undertaken to ascertain whether it interacts directly or indirectly with SnTox3, and also whether it interacts with other host proteins. For example, one possibility was be Pathogenesis-related 1 (PR-1) proteins, as they have been found to interact with SnTox3 (Breen et al., 2016).

Further to this, it would be of great interest to undertake an RNA-seq experiment similar to that conducted in Chapter 3 using a highly SnTox3 sensitive MAGIC founder, such as Rialto. This aim of such an experiment would be to (1) provide gene expression evidence to underpin the bioinformatically predicted genes models for g187 and g188 (including identification and quantification of any alternatively spliced forms), and (2) to investigate the possible gene content in the genomic region between g187 and g188 in a highly sensitive variety. This is because g188 and g187 in these highly sensitivity varieties are 170 Kb apart whereas in Cadenza, an intermediately sensitive variety in the germplasm used for the genome assembly, they are only 1.2 kb apart. Therefore exploring the transcriptome within this interval between these two candidate genes could either provide further evidence in support of g188 and/or g187 or could provide additional key genes that lead to high sensitivity to SnTox3. Indeed, it would be interesting to see whether copy number variation (CNV) for either candidate gene is present in these different wheat genetic backgrounds.

However, whether g188 or g187 is confirmed as *Snn3-B1* is not as important from a breeding perspective. Here, a highly predictive (or perfect) marker for allelic state at *Snn3-B1* is of much greater value. We can see in Chapter 3 that g188 is a clear genetic candidate to provide this perfect marker as, thus far, it has shown perfect clustering between two different versions of this gene (and absence) with the SnTox3 response phenotype in the 19 sequenced wheat genomes investigated (Table 3.5).

As mentioned in Chapter 3, a number of Kompetitive Allele Specific PCR (KASP) markers have been designed for this gene. A good starting place in this venture, would to validate these markers and test them using a panel, such as the AM panel used in Chapter 2, to see if g188 genotype calls continue to be diagnostic of the SnTox3 response phenotype.

Once *Snn3-B1* has been officially identified, further analysis of key differentially expressed genes (DEGs) identified in Chapter 3, could provide a greater idea of how the SnTox3 response pathway functions. For example, it could be interesting to identify the specific role that the three ethylene responsive transcription factors, identified on 1B, play as they were only observed after infiltration with SnTox3 (Figure 3.8). A key question would be whether these have a vital role in SnTox3 triggered hypersensitive response and if so, are they also involved in SnToxA and SnTox1 triggered hypersensitive response.

In Chapter 2, nine minor SnTox3 response quantitative trait loci (QTL) were identified, all of which are novel (Tables 2.1, 2.2, 2.3). To further understanding of the role these QTL play, near isogenic lines (NILs) could be developed, from MAGIC or Avalon × Cadenza lines, to (1) mendalise the QTL, allowing QTL effect to be confirmed and further characterised at the phenotypic and gene expression level, and (2) genetically map the QTL more finely and identify candidate genes for further analysis. Some of which already have some candidates identified in Chapter 3 as DEGs were present in the interval. At the same time, it would be worthwhile to create NILs and linked markers for *QSnb.niab-5A.2* as this was identified as a major QTL for both leaf and glume blotch, the two components of Stagonospora Nodorum Blotch (SNB) in the adult crop (Table 4.4).

Furthermore, the Avalon × Cadenza mapping population has been a key component throughout these studies, linking all three chapters. One approach to fine-mapping a number of the SnTox3 sensitivity QTL as well as the field QTL, without the need to further field or glasshouse work, would be to enrich the A×C map with single nucleotide polymorphism (SNP) data from the 820K Axiom® wheat array (Winfield et al., 2016). This would allow for greater marker density and therefore better genetic resolution. Although ultimately, resolution will be limited by the number of lines investigated.

5.2 Conclusion and Impact of Study

The work carried out here has added to the wealth of SNB knowledge in a number of ways. Firstly, two very promising candidates for *Snn3-B1* were identified, one of which is particularly promising based on differential gene expression between Avalon and Cadenza and with perfect phenotype genotype correlation in the 19

wheat varieties with genome assemblies. A marker set designed to be capable of calling the three major g188 alleles identified, consisting of two KASP markers, has been designed and awaits validation in UK or European panel. It is hoped that they will provide a highly diagnostic (or even perfect) marker set for allelic state at *Snn3-B1*. These markers are likely to be of use to breeders given that avirulent *P. nodorum* strains become virulent with the addition of SnTox3, and *Snn3-B1* has been found to play a role in field resistance to *P. nodorum* (Waters et al., 2011; Ruud et al., 2017). Since the initial submission of this thesis, these KASP markers have been run on a small selection of Australian commercial lines by H. Phan (CCDM, Curtin University), providing promising results in determined high or low sensitivity levels (D.1). However, further validation and adjustment may be needed to use these KASP markers in commercial breeding, particularly in order to identify intermediate sensitivity.

The *P. nodorum* field resistance study conducted in Chapter 4 identified a major QTL on 5A chromosome for both leaf and glume blotch. Given that there have been at least 13 SNB field studies conducted to date with at least 86 QTL discovered, this is of particular value as QTL for both leaf and glume blotch remain rare. As the 5A QTL was discovered using a SNP array, this study is able to provide a physical location for the QTL so that should this QTL be found again in other studies, or explored further in the A×C germplasm, the data is will be easily usable and comparable.

Furthermore, these studies also provide nine QTL for SnTox3 sensitivity, bringing the total number of bread wheat SnTox3 sensitivity QTL to twelve. This is far more than that found for sensitivity to any other effector (Table 4.2). Further exploration of these minor QTL may facilitate further understanding of the effector response pathway for SnTox3, and potentially for other wheat necrotrophic pathogen effectors.

Moreover, these studies have provided a number of genetic resources. Firstly, the new A×C genetic map, which is denser than its predecessors and will made available to use. Secondly, Chapter 3 is first transcriptome study to explore the first 24 hours after infiltration with SnTox3 and provides us with the first insights into pathways that it is triggering after initial exposure. This dataset will be added to the publicly available Wheat Expression Browser, as presently no dataset relating to SNB is included, and the reads will be added to the NCBI database.

5.3 Concluding Remarks

While SNB is now being removed from wheat UK Recommended List requirements for future years, it still remains a significant threat to wheat production globally. Indeed, SNB is still a major problem in Australia, Scandinavia and the USA, and is

also moving into new territory, such as in India. Therefore the data and analysis produced here will be relevant to breeding the genetic resistance needed to tackle this disease on a global scale.

5.4 Work Relating to this Thesis

In addition to the work presented in this thesis, an extensive SnToxA, SnTox1 and SnTox3 screen was carried out on the BreedWheat panel as part of the FSOV funded Wheat effector assisted breeding for resistance to fungal pathogens (WEAB) project. As well as a RNA-seq screen, data, acquired from the NCBI Short Read Archive, was aligned to the *PtrToxB* QTL region to inform candidate gene selection. This was included as part of the Corsi et al. paper which currently under review. A copy of the submitted manuscript can be found in Appendix 4.

Bibliography

Abeysekara, N. S., J. D. Faris, S. Chao, P. E. McClean and T. L. Friesen (2011). Whole-Genome QTL Analysis of Stagonospora nodorum Blotch Resistance and Validation of the SnTox4–*Snn4* Interaction in Hexaploid Wheat. *Phytopathology* **102,** 94–104.

Afzal, A. J., A. J. Wood and D. A. Lightfoot (2008). Plant receptor-like serine threonine kinases: roles in signaling and plant defense. *Molecular Plant-Microbe Interactions* **21,** 507–517.

Aguilar, V., P. Stamp, M. Winzeler, H. Winzeler, G. Schachermayr, B. Keller, S. Zanetti and M. M. Messmer (2005). Inheritance of field resistance to *Stagonospora nodorum* leaf and glume blotch and correlations with other morphological traits in hexaploid wheat (*Triticum aestivum* L.) *Theoretical and Applied Genetics* **111,** 325–336.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215,** 403–10.

Amalraj, A., S. Luang, M. Y. Kumar, P. Sornaraj, O. Eini, N. Kovalchuk, N. Bazanova, Y. Li, N. Yang, S. Eliby et al. (2016). Change of function of the wheat stress-responsive transcriptional repressor Ta RAP 2.1 L by repressor motif modification. *Plant biotechnology journal* **14,** 820–832.

Anders, S. and W. Huber (2010). Differential expression analysis for sequence count data. *Genome Biology* **11,** R106.

Appels, R., K. Eversole, C. Feuillet, B. Keller, J. Rogers, N. Stein, C. J. Pozniak, N. Stein, F. Choulet, A. Distelfeld et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361,** 1–13.

Araus, J. L., G. A. Slafer, C. Royo and M. D. Serret (2008). Breeding for yield potential and stress adaptation in cereals. *Critical Reviews in Plant Sciences* **27,** 377–412.

Badaeva, E. D., O. S. Dedkova, G. Gay, V. A. Pukhalskyi, A. V. Zelenin, S. Bernard and M. Bernard (2007). Chromosomal rearrangements in wheat: their types and distribution. *Genome* **50,** 907–926.

Bahat, A., I. Gelernter, M. B. Brown and Z. Eyal (1980). Factors Affecting the Vertical Progression of Septoria Leaf Blotch in Short-Statured Wheats. *Phytopathology* **70,** 179–184.

Ballance, G. M., L. Lamari and C. C. Bernier (1989). Purification and characterization of a host-selective necrosis toxin from *Pyrenophora tritici-repentis*. *Physiological and Molecular Plant Pathology* **35,** 203–213.

Bates, D., M. Mächler, B. M. Bolker and S. C. Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67,** 1–48.

Berrocal-Lobo, M. and A. Molina (2004). Ethylene response factor 1 mediates Arabidopsis resistance to the soilborne fungus Fusarium oxysporum. *Molecular Plant-microbe Interactions* **17,** 763–770.

Bhathal, J. S., R. Loughman and J. Speijers (2003). Yield reduction in wheat in relation to leaf disease from yellow (tan) spot and septoria nodorum blotch. *European Journal of Plant Pathology* **109,** 435–443.

Blixt, E., A. Djurle, J. Yuen and Å. Olson (2009). Fungicide sensitivity in Swedish isolates of *Phaeosphaeria nodorum*. *Plant Pathology* **58,** 655–664.

Bolser, D. M., D. M. Staines, E. Perry and P. J. Kersey (2017). Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomic data. *Plant Genomics Databases*. Springer, 1–31.

Borrill, P., N. Adamski and C. Uauy (2015). Genomics as the key to unlocking the polyploid potential of wheat. *New Phytologist* **208,** 1008–1022.

Braun, D. M. and J. C. Walker (1996). Plant transmembrane receptors: new pieces in the signaling puzzle. *Trends in biochemical sciences* **21,** 70–73.

Bray, N., H. Pimentel, P. Melsted and L. Pachter (2016). Near-optimal RNA-Seq quantification with kallisto. *Nature Biotechnology* **34,** 525–527.

Breen, S., S. J. Williams, M. Outram, B. Kobe and P. S. Solomon (2017). Emerging Insights into the Functions of Pathogenesis-Related Protein 1. *Trends in Plant Science* **22,** 871–879.

Breen, S., S. J. Williams, B. Winterberg, B. Kobe and P. S. Solomon (2016). Wheat PR-1 proteins are targeted by necrotrophic pathogen effector proteins. *Plant Journal* **88,** 13–25.

Brennan, R. M., B. D. L. Fitt, G. S. Taylor and J. Colhoun (1985). Dispersal of *Septoria nodorum* pycnidiospores by simulated rain and wind. *Journal of Phytopathology* **112,** 291–297.

Broman, K. W., H. Wu, Ś. Sen and G. A. Churchill (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19,** 889–890.

Brueggeman, R., N. Rostoks, D. Kudrna, A. Kilian, F. Han, J. Chen, A. Druka, B. Steffenson and A. Kleinhofs (2002). The barley stem rust-resistance gene Rpg1

is a novel disease-resistance gene with homology to receptor kinases. *Proceedings of the National Academy of Sciences* **99,** 9328–9333.

Campbell, I. D. and P. Bork (1993). Epidermal growth factor-like modules. *Current Opinion in Structural Biology* **3,** 385–392.

Cat, A., M. Tekin, T. Akar and M. Catal (2018). First report of Stagonospora nodorum blotch caused by *Parastagonospora nodorum* on emmer wheat (*Triticum dicoccum* Schrank) in Turkey. *Journal of Plant Pathology* **17,** 42161.

Chowdhury, S., A. Basu and S. Kundu (2017). Biotrophy-necrotrophy switch in pathogen evoke differential response in resistant and susceptible sesame involving multiple signaling pathways at different phases. *Scientific reports* **7,** 1–17.

Chu, C. G., J. D. Faris, S. S. Xu and T. L. Friesen (2010). Genetic analysis of disease susceptibility contributed by the compatible *Tsn1*-SnToxA and *Snn1*-SnTox1 interactions in the wheat-*Stagonospora nodorum* pathosystem. *Theoretical and Applied Genetics* **120,** 1451–1459.

Clavijo, B. J., L. Venturini, C. Schudoma, G. G. Accinelli, G. Kaithakottil, J. Wright, P. Borrill, G. Kettleborough, D. Heavens, H. Chapman et al. (2017). An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Research* **27,** 885–896.

Cockram, J., H. Jones, F. J. Leigh, D. O'Sullivan, W. Powell, D. A. Laurie and A. J. Greenland (2007). Control of flowering time in temperate cereals: Genes, domestication, and sustainable productivity. *Journal of Experimental Botany* **58,** 1231–1244.

Cockram, J. and I. J. Mackay (2018). Genetic mapping populations for conducting high resolution trait mapping in plants. *Advances in Biochemical Engineering Biotechnology*. Varshney, Berlin: Springer International Publishing.

Cockram, J., A. Scuderi, T. Barber, E. Furuki, K. A. Gardner, N. Gosman, R. Kowalczyk, H. Phan, G. A. Rose, K.-C. Tan et al. (2015). Fine-mapping the wheat *Snn1* locus conferring sensitivity to the *Parastagonospora nodorum* necrotrophic effector SnTox1 using an eight founder multiparent advanced generation inter-cross population. *G3* **5,** 2257–2266.

Corsi, B., L. Percival-Alwyn, R. Downie, L. Venturini, E. Iagallo, C. Mantello, C. McCormick-Barnes, P. See, R. Oliver, C. Moffat et al. (2020). Genetic analysis of wheat sensitivity to the ToxB fungal effector from *Pyrenophora tritici-repentis*, the causal agent of tan spot. *Theoretical and Applied Genetics* **133,** 935–950.

Cui, F., X. Fan, C. Zhao, W. Zhang, M. Chen, J. Ji and J. Li (2014). A novel genetic map of wheat : Utility for mapping QTL for yield under different nitrogen treatments A novel genetic map of wheat : utility for mapping QTL for yield under different nitrogen treatments. *BMC Genetics* **15,** 1–17.

Czembor, P. C., E. Arseniuk, A. Czaplicki, Q. Song, P. B. Cregan and P. P. Ueng (2003). QTL mapping of partial resistance in winter wheat to *Stagonospora nodorum* blotch. *Genome* **46,** 546–554.

De Wit, P. J. G. M., R. Mehrabi, H. A. Van Den Burg and I. Stergiopoulos (2009). Fungal effector proteins: Past, present and future: Review. *Molecular Plant Pathology* **10,** 735–747.

Dodds, P. N. and J. P. Rathjen (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. *Nature Reviews Genetics* **11,** 539.

Dong, Y., Q. Zhao, X. Liu, X. Zhang, Z. Qi, H. Zhang, X. Zheng and Z. Zhang (2015). MoMyb1 is required for asexual development and tissue-specific infection in the rice blast fungus *Magnaporthe oryzae*. *BMC microbiology* **15,** 37.

Downie, R., M. Lin, B. Corsi, A. Ficke, M. Lillemo, R. Oliver, H. Phan, T. KC and J. Cockram (submitted). Septoria nodorum blotch of wheat: disease management and resistance breeding in the face of shifting disease dynamics and a changing environment.

Downie, R. C., L. Bouvet, E. Furuki, N. Gosman, K. A. Gardner, I. J. Mackay, C. Campos Mantello, G. Mellers, H. T. T. Phan, G. A. Rose et al. (2018). Assessing European Wheat Sensitivities to *Parastagonospora nodorum* Necrotrophic Effectors and Fine-Mapping the *Snn3-B1* Locus Conferring Sensitivity to the Effector SnTox3. *Frontiers in Plant Science* **9.**

Duba, A., K. Goriewa-Duba and U. Wachowska (2018). A review of the interactions between wheat and wheat pathogens: *Zymoseptoria tritici*, *Fusarium* spp. and *Parastagonospora nodorum*. *International journal of molecular sciences* **19,** 1138.

Dufresne, M. and A. E. Osbourn (2001). Definition of tissue-specific and general requirements for plant infection in a phytopathogenic fungus. *Molecular plant-microbe interactions* **14,** 300–307.

Dzhavakhiya, V., L. Shcherbakova, Y. Semina, N. Zhemchuzhina and B. Campbell (2012). Chemosensitization of plant pathogenic fungi to agricultural fungicides. *Frontiers in Microbiology* **3,** 1–9.

Eyal, Z. (1981). Integrated Control of Septoria Diseases of Wheat. *Plant Disease* **65,** 763–768.

Eyal, Z. (1999). The septoria tritici and stagonospora nodorum blotch diseases of wheat. *European Journal of Plant Pathology* **105,** 629–641.

FAO (2009). *Global agriculture towards 2050.*

FAO (2015). *Cereal Supply and Demand Brief.*

Faris, J. D., J. A. Anderson, L. J. Francl and J. G. Jordahl (1996). Chromosomal location of a gene conditioning insensitivity in wheat to a necrosis-inducing culture filtrate from *Pyrenophora tritici-repentis*. *Phytopathology* **86,** 459–463.

Faris, J. D., Z. Zhang, H. Lu, S. Lu, L. Reddy, S. Cloutier, J. P. Fellers, S. W. Meinhardt, J. B. Rasmussen, S. S. Xu et al. (2010). A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens. *Proceedings of the National Academy of Sciences* **107,** 13544–13549.

Faris, J. D., Z. Zhang, J. B. Rasmussen and T. L. Friesen (2011). Variable Expression of the Stagonospora nodorum Effector SnToxA Among Isolates Is Correlated with Levels of Disease in Wheat. *Molecular Plant-Microbe Interactions* **24,** 1419–1426.

Finn, R. D., P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas et al. (2016). The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Research* **44,** D279–D285.

Francki, M. G., M. Shankar, E. Walker, R. Loughman, H. Golzar and H. Ohm (2011). New Quantitative Trait Loci in Wheat for Flag Leaf Resistance to Stagonospora nodorum Blotch. *Phytopathology* **101,** 1278–1284.

Friesen, T. L., S. Ali, K. K. Klein and J. B. Rasmussen (2005). Population genetic analysis of a global collection of *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. *Phytopathology* **95,** 1144–50.

Friesen, T. L., D. J. Holmes, R. L. Bowden and J. D. Faris (2018). ToxA is present in the U.S. *Bipolaris sorokiniana* population and is a significant virulence factor on wheat harboring *Tsn1*. *Plant Disease* **102,** 2446–2452.

Friesen, T. L., Z. Zhang, P. S. Solomon, R. P. Oliver and F. J. D (2008). Characterization of the interaction of a novel *Stagonospora nodorum* host-selective toxin with a wheat susceptibility gene. *Plant Physiology* **146,** 682–693.

Friesen, T. L., C. G. Chu, Z. H. Liu, S. S. Xu, S. Halley and F. J. D (2009). Host-selective toxins produced by *Stagonospora nodorum* confer disease susceptibility in adult wheat plants under field conditions. *Theoretical and Applied Genetics* **118,** 1489–1497.

Friesen, T. L., C. Chu, S. S. Xu and J. D. Faris (2012). SnTox5-*Snn5*: A novel *Stagnospora nodorum* effector-wheat gene interaction and its relationship with the SnToxA-*Tsn1* and SnTox3-*Snn3* interactions. *Molecular Plant Pathology* **13,** 1101–1109.

Friesen, T. L., S. W. Meinhardt and J. D. Faris (2007). The *Stagonospora nodorum*-wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. *The Plant Journal* **51,** 681–692.

Friesen, T. L., E. H. Stukenbrock, Z. Liu, S. Meinhardt, H. Ling, J. D. Faris, J. B. Rasmussen, P. S. Solomon, B. A. McDonald and R. P. Oliver (2006). Emergence of

a new disease as a result of interspecific virulence gene transfer. *Nature Genetics* **38,** 953–956.

Fulton, T. M., J. Chunwongse and S. D. Tanksley (1995). Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter* **13,** 207–209.

Gao, Y., J. D. Faris, Z. Liu, Y. M. Kim, R. A. Syme, R. P. Oliver, S. S. Xu and T. L. Friesen (2015). Identification and Characterization of the SnTox6-*Snn6* Interaction in the *Parastagonospora nodorum*–Wheat Pathosystem. *Molecular Plant-Microbe Interactions* **28,** 615–625.

Gardner, K. A., L. M. Wittern and I. J. Mackay (2016). A highly recombined, high-density, eight-founder wheat MAGIC map reveals extensive segregation distortion and genomic locations of introgression segments. *Plant Biotechnology Journal* **14,** 1406–1417.

González-Lamothe, R., D. I. Tsitsigiannis, A. A. Ludwig, M. Panicot, K. Shirasu and J. D. G. Jones (2006). The U-box protein CMPG1 is required for efficient activation of defense mechanisms triggered by multiple resistance genes in tobacco and tomato. *The Plant Cell* **18,** 1067–83.

Guo, Y.-H., Y.-P. Yu, D. Wang, C.-A. Wu, G.-D. Yang, J.-G. Huang and C.-C. Zheng (2009). GhZFP1, a novel CCCH-type zinc finger protein from cotton, enhances salt stress tolerance and fungal disease resistance in transgenic tobacco by interacting with GZIRD21A and GZIPR5. *New Phytologist* **183,** 62–75.

Hane, J. K., R. G. Lowe, P. S. Solomon, K.-C. Tan, C. L. Schoch, J. W. Spatafora, P. W. Crous, C. Kodira, B. W. Birren, J. E. Galagan et al. (2007). Dothideomycete Plant Interactions Illuminated by Genome Sequencing and EST Analysis of the Wheat Pathogen *Stagonospora nodorum*. *The Plant Cell* **19,** 3347–3368.

Hane, J. K., J. Paxman, D. A. Jones, R. P. Oliver and P. de Wit (2020). "CATAStrophy," a Genome-Informed Trophic Classification of Filamentous Plant Pathogens–How Many Different Types of Filamentous Plant Pathogens Are There? *Frontiers in microbiology* **10,** 3088.

Hassani-Pak, K., M. Castellote, M. Esch, M. Hindle, A. Lysenko, J. Taubert and C. Rawlings (2016). Developing integrated crop knowledge networks to advance candidate gene discovery. *Applied & Translational Genomics* **11,** 18–26.

Heitz, T., E. Widemann, R. Lugan, L. Miesch, P. Ullmann, L. Désaubry, E. Holder, B. Grausem, S. Kandel, M. Miesch et al. (2012). Cytochromes P450 CYP94C1 and CYP94B3 catalyze two successive oxidation steps of plant hormone jasmonoyl-isoleucine for catabolic turnover. *Journal of Biological Chemistry* **287,** 6296–6306.

Hong, J. K., H. W. Choi, I. S. Hwang and B. K. Hwang (2007). Role of a novel pathogen-induced pepper C3–H–C4 type RING-finger protein gene, CaRFP1, in

disease susceptibility and osmotic stress tolerance. *Plant molecular biology* **63,** 571–588.

Huang, B. E. and A. W. George (2011). R/mpMap: A computational platform for the genetic analysis of multiparent recombinant inbred lines. *Bioinformatics* **27,** 727–729.

IRRI (2007). *CROPSTAT for windows, Version 7.2.* Metro Manila, Philippines.

IWGSC et al. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* **345,** 1251788.

Jighly, A., M. Alagu, F. Makdis, M. Singh, S. Singh, L. C. Emebiri and F. C. Ogbonnaya (2016). Genomic regions conferring resistance to multiple fungal pathogens in synthetic hexaploid wheat. *Molecular Breeding* **36,** 1–19.

Joazeiro, C. A. and A. M. Weissman (2000). RING finger proteins: mediators of ubiquitin ligase activity. *Cell* **102,** 549–552.

Jones, J. D. and J. L. Dangl (2006). The plant immune system. *nature* **444,** 323.

Kang, H. M., N. A. Zaitlen, C. M. Wade, A. Kirby, D. Heckerman, M. J. Daly and E. Eskin (2008). Efficient control of population structure in model organism association mapping. *Genetics* **178,** 1709–1728.

Katoch, S., S. K. Rana and P. N. Sharma (2018). Application of PCR based diagnostics in the exploration of *Parastagonospora nodorum* prevalence in wheat growing regions of Himachal Pradesh. *Journal of Plant Biochemistry and Biotechnology* **7.**

Kim, D., B. Langmead and S. Salzberg (2017). *HISAT2: graph-based alignment of next-generation sequencing reads to a population of genomes.*

Kishor, P. K., S. Sangam, R. Amrutha, P. S. Laxmi, K. Naidu, K. Rao, S. Rao, K. Reddy, P. Theriappan and N. Sreenivasulu (2005). Regulation of proline biosynthesis, degradation, uptake and transport in higher plants: its implications in plant growth and abiotic stress tolerance. *Current Science* **88,** 424–438.

Kliebenstein, D. J. and H. C. Rowe (2008). Ecological costs of biotrophic versus necrotrophic pathogen resistance, the hypersensitive response and signal transduction. *Plant Science* **174,** 551–556.

Krapp, A. R., V. B. Tognetti, N. Carrillo and A. Acevedo (1997). The Role of Ferredoxin-NADP+ Reductase in the Concerted Cell Defense Against Oxidative Damage: Studies using Escherichia Coli Mutants and Cloned Plant Genes. *European journal of biochemistry* **249,** 556–563.

Krasileva, K. V., V. Buffalo, P. Bailey, S. Pearce, S. Ayling, F. Tabbita, M. Soria, S. Wang, E. Akhunov, C. Uauy et al. (2013). Separating homeologs by phasing in the tetraploid wheat transcriptome. *Genome biology* **14,** R66.

Krasileva, K. V., H. A. Vasquez-Gross, T. Howell, P. Bailey, F. Paraiso, L. Clissold, J. Simmonds, R. H. Ramirez-Gonzalez, X. Wang, P. Borrill et al. (2017). Uncovering

hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences* **114,** E913–E921.

Krueger, F. (2012). *Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type libraries.*

Larroque, M., E. Belmas, T. Martinez, S. Vergnes, N. Ladouce, C. Lafitte, E. Gaulin and B. Dumas (2013). Pathogen-associated molecular pattern-triggered immunity and resistance to the root pathogen Phytophthora parasitica in Arabidopsis. *Journal of experimental botany* **64,** 3615–3625.

Lawrence-Dill, C. (2019). *GOMAP Wheat Reference Sequences 1.1.*

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34,** 3094–3100.

Li, H. and R. Durbin (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26,** 589–595.

Li, P., J. G. Huang, S. W. Yu, Y. Y. Li, P. Sun, C. A. Wu and C. C. Zheng (2016). *Arabidopsis* YL1/BPG2 is Involved in Seedling Shoot Response to Salt Stress through ABI4. *Scientific Reports* **6,** 1–11.

Li, S., J. Jia, X. Wei, X. Zhang, L. Li, Y. Xu, F. Jiang, H. Wang and L. Li (2007). A intervarietal genetic map and QTL analysis for yield traits in wheat. *Molecular Breeding* **20,** 167–178.

Lin, M., B. Corsi, A. Ficke, K.-C. Tan, J. Cockram and L. M (2020). Genetic mapping using a wheat multi-founder population reveals a locus on chromosome 2A controlling resistance to both leaf and glume blotch caused by the necrotrophic fungal pathogen *Parastagonospora nodorum*. *Theoretical and Applied Genetics* **133,** 785–808.

Lipka, A. E., F. Tian, Q. Wang, J. Peiffer, M. Li, P. J. Bradbury, M. A. Gore, E. S. Buckler and Z. Zhang (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics* **28,** 2397–2399.

Liu, Y. and C. He (2017). A review of redox signaling and the control of MAP kinase pathway in plants. *Redox biology* **11,** 192–204.

Liu, Z., J. D. Faris, S. W. Meinhardt, S. Ali, J. B. Rasmussen and T. L. Friesen (2004). Genetic and physical mapping of a gene conditioning sensitivity in wheat to a partially purified host-selective toxin produced by *Stagonospora nodorum*. *Phytopathology* **94,** 1056–1060.

Liu, Z., J. D. Faris, R. P. Oliver, K. C. Tan, P. S. Solomon, M. C. McDonald, B. A. McDonald, A. Nunez, S. Lu, J. B. Rasmussen et al. (2009). SnTox3 acts in effector triggered susceptibility to induce disease on wheat carrying the *Snn3* gene. *PLoS Pathogens* **5.**

Liu, Z., T. L. Friesen, H. Ling, S. W. Meinhardt, R. P. Oliver, J. B. Rasmussen and J. D. Faris (2006). The *Tsn1*-ToxA interaction in the wheat-*Stagonospora nodorum* pathosystem parallels that of the wheat-tan spot system. *Genome* **49,** 1265–1273.

Liu, Z., Z. Zhang, J. D. Faris, R. P. Oliver, R. Syme, M. C. McDonald, B. A. McDonald, P. S. Solomon, S. Lu, W. L. Shelver et al. (2012). The cysteine rich necrotrophic effector SnTox1 produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harboring *Snn1*. *PLoS Pathogens* **8,** e1002467.

Loginova, D. and O. Silkova (2018). The genome of bread wheat *Triticum aestivum* L.: unique structural and functional properties. *Russian Journal of Genetics* **54,** 403–414.

Loon, L. C. van, M. Rep and C. M. Pieterse (2006). Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol.* **44,** 135–162.

Lorang, J. M., T. A. Sweat and T. J. Wolpert (2007). Plant disease susceptibility conferred by a "resistance" gene. *Proceedings of the National Academy of Sciences* **104,** 14861–14866.

Lorrain, C., K. C. Gonçalves dos Santos, H. Germain, A. Hecker and S. Duplessis (2019). Advances in understanding obligate biotrophy in rust fungi. *New Phytologist* **222,** 1190–1206.

Lowe, R. G., M. Lord, K. Rybak, R. D. Trengove, R. P. Oliver and P. S. Solomon (2009). Trehalose biosynthesis is involved in sporulation of Stagonospora nodorum. *Fungal Genetics and Biology* **46,** 381–389.

Lu, Q. and M. Lillemo (2014). Molecular mapping of adult plant resistance to *Parastagonospora nodorum* leaf blotch in bread wheat lines 'Shanghai-3/Catbird' and 'Naxos'. *Theoretical and Applied Genetics* **127,** 2635–2644.

Ma, J., L. U. Wingen, S. Orford, P. Fenwick, J. Wang and S. Griffiths (2015). Using the UK reference population Avalon × Cadenza as a platform to compare breeding strategies in elite Western European bread wheat. *Molecular Breeding* **35.**

Mackay, I. J., P. Bansept-Basler, T. Barber, A. R. Bentley, J. Cockram, N. Gosman, A. J. Greenland, R. Horsnell, R. Howells, D. M. O'Sullivan et al. (2014). An eight-parent multiparent advanced generation inter-cross population for winter-sown wheat: creation, properties, and validation. *G3* **4,** 1603–1610.

Madeira, F., J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. Tivey, S. C. Potter, R. D. Finn, R. Lopez et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research* **47,** W636–W641.

Manning, V. A. and L. M. Ciuffetti (2005). Localization of Ptr ToxA Produced by *Pyrenophora tritici-repentis* Reveals Protein Import into Wheat Mesophyll Cells. *The Plant Cell* **17,** 3203–3212.

Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales et al. (2010). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research* **39,** D225–D229.

McCallum, C. M., L. Comai, E. Greene and S. Henikoff (2000). Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiology* **123,** 439–442.

McDonald, B. A. and C. Linde (2002). Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40,** 349–379.

McDonald, B. A., J. Miles, L. Nelson and R. E. Pettway (1994). Genetic variability in nuclear DNA in field populations of *Stagonospora nodorum*. *Phytopathology* **84,** 250–255.

McDonald, M. C., R. P. Oliver, F. T. L, P. C. Brunner and B. A. McDonald (2013). Global diversity and distribution of three necrotrophic effectors in *Phaeosphaeria nodorum* and related species. *New Phytologist* **199,** 241–251.

McIntosh, R. A., Y. Yamazaki, J. Dubcovsky, J. Rogers, C. Morris and D. J. Somers (2008). Catalogue of gene symbols for wheat. *Proceedings of the 11th International Wheat Genetics Symposium.* San Diego, CA, 24–29.

Meng, L., H. Li, L. Zhang and J. Wang (2015). QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *The Crop Journal* **3,** 269–283.

Mengiste, T. (2012). Plant immunity to necrotrophs. *Annual review of phytopathology* **50,** 267–294.

Mitchell, A. L., T. K. Attwood, P. C. Babbitt, M. Blum, P. Bork, A. Bridge, S. D. Brown, H.-Y. Chang, S. El-Gebali, M. I. Fraser et al. (2018). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic acids research* **47,** D351–D360.

Möller, M. and E. H. Stukenbrock (2017). Evolution and genome architecture in fungal plant pathogens. *Nature Reviews Microbiology* **15,** 756.

Murray, G. M. and J. P. Brennan (2009). Estimating disease losses to the Australian wheat industry. *Australasian Plant Pathology* **38,** 558.

Nagy, E. D. and J. L. Bennetzen (2008). Pathogen corruption and site-directed recombination at a plant disease resistance gene cluster. *Genome research* **18,** 1918–1923.

Nawrath, C., S. Heck, N. Parinthawong and J.-P. Métraux (2002). EDS5, an essential component of salicylic acid–dependent signaling for disease resistance in Arabidopsis, is a member of the MATE transporter family. *The Plant Cell* **14,** 275–286.

Nekrasov, V., J. Li, M. Batoux, M. Roux, Z.-H. Chu, S. Lacombe, A. Rougon, P. Bittel, M. Kiss-Papp, D. Chinchilla et al. (2009). Control of the pattern-recognition receptor EFR by an ER protein complex in plant immunity. *The EMBO journal* **28,** 3428–3438.

Nicholson, P., H. N. Rezanoor and A. J. Worland (1993). Chromosomal Location of Resistance to *Septoria nodorum* in a Synthetic Hexaploid Wheat Determined by the Study of Chromosomal Substitution Lines in 'Chinese Spring' Wheat. *Plant Breeding* **110,** 177–184.

Oliver, R. P., T. L. Friesen, J. D. Faris and P. S. Solomon (2012). *Stagonospora nodorum*: from pathology to genomics and host resistance. *Annual Review of Phytopathology* **50,** 23–43.

Oliver, R. P. and S. V. Ipcho (2004). Arabidopsis pathology breathes new life into the necrotrophs-vs.-biotrophs classification of fungal pathogens. *Molecular Plant Pathology* **5,** 347–352.

Oliver, R. P. and P. Solomon (2010). New developments in pathogenicity and virulence of necrotrophs. *Current Opinion in Plant Biology* **13,** 415–419.

Pandelova, I., M. Figueroa, L. J. Wilhelm, V. A. Manning, A. N. Mankaney, T. C. Mockler and L. M. Ciuffetti (2012). Host-selective toxins of Pyrenophora tritici-repentis induce common responses associated with host susceptibility. *PLoS One* **7,** e40240.

Pereira, D. A., B. A. McDonald and P. C. Brunner (2017). Mutations in the CYP51 gene reduce DMI sensitivity in *Parastagonospora nodorum* populations in Europe and China. *Pest Management Science* **73,** 1503–1510.

Petersen, G., O. Seberg, M. Yde and K. Berthelsen (2006). Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). *Molecular phylogenetics and evolution* **39,** 70–82.

Phan, H., K. Rybak, S. Bertazzoni, E. Furuki, E. Dinglasan, L. T. Hickey, R. P. Oliver and K. C. Tan (2018). Novel sources of resistance to Septoria nodorum blotch in the Vavilov wheat collection identified by genome-wide association studies. *Theoretical and Applied Genetics* **131,** 1223–1238.

Phan, H., K. Rybak, E. Furuki, S. Breen, P. S. Solomon, R. P. Oliver and K. C. Tan (2016). Differential effector gene expression underpins epistasis in a plant fungal disease. *Plant Journal* **87,** 343–354.

Pružinská, A., G. Tanner, I. Anders, M. Roca and S. Hörtensteiner (2003). Chlorophyll breakdown: pheophorbide a oxygenase is a Rieske-type iron–sulfur protein, encoded by the accelerated cell death 1 gene. *Proceedings of the National Academy of Sciences* **100,** 15259–15264.

Qi, Y., J. Zhao, R. An, J. Zhang, S. Liang, J. Shao, X. Liu, L. An and F. Yu (2016). Mutations in circularly permuted GTPase family genes *AtNOA1/RIF1/SVR10* and *BPG2* suppress *var2*-mediated leaf variegation in *Arabidopsis thaliana*. *Photosynthesis Research* **127,** 355–367.

Quaedvlieg, W., G. J. M. Verkley, H.-D. Shin, R. W. Barreto, A. C. Alfenas, W. J. Swart, J. Z. Groenewald and P. W. Crous (2013). Sizing up Septoria. *Studies in Mycology* **75,** 307–90.

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna.

Ramirez-Gonzalez, R. H., C. Uauy and M. Caccamo (2015). PolyMarker: A fast polyploid primer design pipeline. *Bioinformatics* **31,** 2038–2039.

Rice, P., I. Longden and A. Bleasby (2000). *EMBOSS: the European molecular biology open software suite*.

Richards, J., G. Ameen and R. Brueggeman (Nov. 2017). Inverse Gene-For-Gene: Necrotrophic specialist's modus operandi in Barley and Wheat, 606–633.

Rinaldo, A., B. Gilbert, R. Boni, S. G. Krattinger, D. Singh, R. F. Park, E. Lagudah and M. Ayliffe (2017). The Lr34 adult plant rust resistance gene provides seedling resistance in durum wheat without senescence. *Plant biotechnology journal* **15,** 894–905.

Ruud, A. K., J. A. Dieseth and M. Lillemo (2018). Effects of three *Parastagonospora nodorum* necrotrophic effectors on spring wheat under Norwegian field conditions. *Crop Science* **58,** 159–168.

Ruud, A. K., S. Windju, T. Belova, T. L. Friesen and M. Lillemo (2017). Mapping of SnTox3–*Snn3* as a major determinant of field susceptibility to Septoria nodorum leaf blotch in the SHA3/CBRD × Naxos population. *Theoretical and Applied Genetics* **130,** 1361–1374.

Saintenac, C., W.-S. Lee, F. Cambon, J. J. Rudd, R. C. King, W. Marande, S. J. Powers, H. Bergès, A. L. Phillips, C. Uauy et al. (2018). Wheat receptor-kinase-like protein Stb6 controls gene-for-gene resistance to fungal pathogen Zymoseptoria tritici. *Nature genetics* **50,** 368.

Salam, K. P., G. J. Thomas, C. Beard, R. Loughman, W. J. MacLeod and M. U. Salam (2013). Application of meta-analysis in plant pathology: A case study examining the impact of fungicides on wheat yield loss from the yellow spot-septoria nodorum blotch disease complex in Western Australia. *Food Security* **5,** 319–325.

Salamov, A. A. and V. V. Solovyev (2000). Ab initio gene finding in Drosophila genomic DNA. *Genome research* **10,** 516–522.

Schnurbusch, T., S. Paillard, D. Fossati, M. Messmer, G. Schachermayr, M. Winzeler and B. Keller (2003). Detection of QTLs for Stagonospora glume blotch resistance in Swiss winter wheat. *Theoretical and Applied Genetics* **107,** 1226–1234.

Scott, P., P. Benedikz and C. J. Cox (1982). A genetic study of the relationship between height, time of ear emergence and resistance to *Septoria nodorum* in wheat. *Plant Pathology* **31,** 45–60.

Shankar, M., E. Walker, H. Golzar, R. Loughman, R. E. Wilson and M. G. Francki (2008). Quantitative Trait Loci for Seedling and Adult Plant Resistance to *Stagonospora nodorum* in Wheat. *Phytopathology* **98,** 886–893.

Shatalina, M., M. Messmer, C. Feuillet, F. Mascher, E. Paux, F. Choulet, T. Wicker and B. Keller (2014). High-resolution analysis of a QTL for resistance to Stagonospora nodorum glume blotch in wheat reveals presence of two distinct resistance loci in the target interval. *Theoretical and Applied Genetics* **127,** 573–586.

Shaw, M. W., S. J. Bearchell, B. D. L. Fitt and B. A. Fraaije (2008). Long-term relationships between environment and abundance in wheat of *Phaeosphaeria nodorum* and *Mycosphaerella graminicola*. *New Phytologist* **177,** 229–238.

Shewry, P. R. (2009). Wheat. *Journal of Experimental Botany* **60,** 1537–1553.

Shi, G., T. L. Friesen, J. Saini, S. S. Xu, J. B. Rasmussen and J. D. Faris (2015). The Wheat Gene Confers Susceptibility on Recognition of the Necrotrophic Effector SnTox7. *The Plant Genome* **8,** 1–10.

Shi, G., Z. Zhang, T. L. Friesen, U. Bansal, S. Cloutier, T. Wicker, J. B. Rasmussen and J. D. Faris (2016a). Marker development, saturation mapping, and high-resolution mapping of the Septoria nodorum blotch susceptibility gene *Snn3-B1* in wheat. *Molecular Genetics and Genomics* **291,** 107–119.

Shi, G., Z. Zhang, T. L. Friesen, D. Raats, T. Fahima, R. S. Brueggeman, S. Lu, H. N. Trick, Z. Liu, W. Chao et al. (2016b). The hijacking of a receptor kinase – driven pathway by a wheat fungal pathogen leads to disease. *Science Advances* **2,** e1600822.

Silva-Gomes, S., A. Decout and J. Nigou (2016). Pathogen-associated molecular patterns (PAMPs). *Compendium of Inflammatory Diseases,* 1055–1069.

Smith, H. E. and S. Ward (1998). Identification of protein-protein interactions of the major sperm protein (MSP) of *Caenorhabditis elegans*. *Journal of molecular biology* **279,** 605–619.

Solomon, P., R. G. T. Lowe, K.-C. Tan, O. D. C. Waters and R. P. Oliver (2006). Stagonospora nodorum: cause of stagonospora nodorum blotch of wheat. *Molecular Plant Pathology* **7,** 147–56.

Sommerhalder, R. J., B. A. McDonald, F. Mascher and J. Zhan (2011). Effect of hosts on competition among clones and evidence of differential selection between

pathogenic and saprophytic phases in experimental populations of the wheat pathogen *Phaeosphaeria nodorum*. *BMC Evolutionary Biology* **11.**

Stanke, M., M. Diekhans, R. Baertsch and D. Haussler (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24,** 637–644.

Steuernagel, B., K. Witek, S. G. Krattinger, R. H. Ramirez-Gonzalez, H.-j. Schoonbeek, G. Yu, E. Baggs, A. Witek, I. Yadav, K. V. Krasileva et al. (2018). Physical and transcriptional organisation of the bread wheat intracellular immune receptor repertoire.

Subramanyam, S., C. Zheng, J. T. Shukle and C. E. Williams (2013). Hessian fly larval attack triggers elevated expression of disease resistance dirigent-like protein-encoding gene, HfrDrd, in resistant wheat. *Arthropod-Plant Interactions* **7,** 389–402.

Tan, K.-C., M. Ferguson-Hunt, K. Rybak, O. D. C. Waters, W. A. Stanley, C. S. Bond, E. H. Stukenbrock, T. L. Friesen, J. D. Faris, B. A. McDonald et al. (2012). Quantitative Variation in Effector Activity of ToxA Isoforms from *Stagonospora nodorum* and *Pyrenophora tritici-repentis*. *Molecular Plant-Microbe Interactions* **25,** 515–522.

Tan, K.-C., O. D. C. Waters, K. Rybak, E. Antoni, E. Furuki and R. P. Oliver (2014). Sensitivity to three *Parastagonospora nodorum* necrotrophic effectors in current Australian wheat cultivars and the presence of further fungal effectors. *Crop and Pasture Science* **65,** 150.

Tanaka, S. (1933). Studies on black pot disease of the Japenese Pear (*Pirus serotina Rehd*). *Memoirs of the College of Agriculture*. Kyoto: Kyoto University, 31.

Tang, C., Q. Xu, M. Zhao, X. Wang and Z. Kang (2018). Understanding the lifestyles and pathogenicity mechanisms of obligate biotrophic fungi in wheat: the emerging genomics era. *The Crop Journal* **6,** 60–67.

Thomma, B. P., T. Nürnberger and M. H. Joosten (2011). Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *The Plant Cell* **23,** 4–15.

Tomas, A., G. H. Feng, G. R. Reeck, W. W. Bockus and J. E. Leach (1990). Purification of a cultivar-specific toxin from *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. *Molecular Plant-Microbe Interactions* **3,** 221–224.

Trujillo, M., K. Ichimura, C. Casais and K. Shirasu (2008). Negative regulation of PAMP-triggered immunity by an E3 ubiquitin ligase triplet in Arabidopsis. *Current Biology* **18,** 1396–1401.

Tsuda, K. and F. Katagiri (2010). Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. *Current opinion in plant biology* **13,** 459–465.

Tuinstra, M. R., G. Ejeta and P. B. Goldsbrough (1997). Heterogeneous inbred family (HIF) analysis: A method for developing near-isogenic lines that differ at quantitative trait loci. *Theoretical and Applied Genetics* **95,** 1005–1011.

UN (2013). *World population projected to reach 9.6 billion by 2050.*

Uphaus, J., E. Walker, M. Shankar, H. Golzar, R. Loughman, M. Francki and H. Ohm (2007). Quantitative trait loci identified for resistance to stagonospora glume blotch in wheat in the USA and Australia. *Crop Science* **47,** 1813–1822.

Vanholme, R., B. Demedts, K. Morreel, J. Ralph and W. Boerjan (2010). Lignin biosynthesis and structure. *Plant Physiology* **153,** 895–905.

Vincent, D., L. A. Du Fall, A. Livk, U. Mathesius, R. J. Lipscombe, R. P. Oliver, T. L. Friesen and P. S. Solomon (2012). A functional genomics approach to dissect the mode of action of the *Stagonospora nodorum* effector protein SnToxA in wheat. *Molecular Plant Pathology* **13,** 467–482.

VSN International (2011). *GenStat for Windows.* Hemel Hempstead.

Wang, S., D. Wong, K. Forrest, A. Allen, S. Chao, B. E. Huang, M. Maccaferri, S. Salvi, S. G. Milner, L. Cattivelli et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal* **12,** 787–796.

Waters, O. D. C., J. Lichtenzveig, K. Rybak, T. L. Friesen and R. P. Oliver (2011). Prevalence and importance of sensitivity to the *Stagonospora nodorum* necrotrophic effector SnTox3 in current Western Australian wheat cultivars. *Crop and Pasture Science* **62,** 556–562.

Wehrens, R. and J. Kruisselbrink (2018). Flexible Self-Organizing Maps in kohonen 3.0. *Journal of Statistical Software* **87,** 1–18.

West, J. S., J. A. Townsend, M. Stevens and B. D. Fitt (2012). Comparative biology of different plant pathogens to estimate effects of climate change on crop diseases in Europe. *European Journal of Plant Pathology* **133,** 315–331.

Wilkinson, P. A., M. O. Winfield, G. L. Barker, S. Tyrrell, X. Bian, A. M. Allen, A. Burridge, J. A. Coghill, C. Waterfall, M. Caccamo et al. (2016). CerealsDB 3.0: Expansion of resources and data integration. *BMC Bioinformatics* **17,** 1–9.

Winfield, M. O., A. M. Allen, A. J. Burridge, G. L. Barker, H. R. Benbow, P. A. Wilkinson, J. Coghill, C. Waterfall, A. Davassi, G. Scopes et al. (2016). High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnology Journal* **14,** 1195–1206.

Winterberg, B., L. A. Du Fall, X. Song, D. Pascovici, N. Care, M. Molloy, S. Ohms and P. S. Solomon (2014). The necrotrophic effector protein SnTox3 re-programs metabolism and elicits a strong defence response in susceptible wheat leaves. *BMC plant biology* **14,** 215.

Xia, Y. (2004). Proteases in pathogenesis and plant defence. *Cellular Microbiology* **6,** 905–913.

Zadoks, J. C., T. T. Chang and C. F. Konzak (1974). A decimal code for the growth stages of cereals. *Weed Research* **14,** 415–421.

Zeng, L.-R., S. Qu, A. Bordeos, C. Yang, M. Baraoidan, H. Yan, Q. Xie, B. H. Nahm, H. Leung and G.-l. Wang (2004). *Spotted leaf11*, a negative regulator of plant cell death and defense, encodes a U-Box/armadillo repeat protein endowed with E3 ubiquitin ligase activity. *The Plant Cell* **16,** 2795–2808.

Zhang, Z., T. L. Friesen, S. S. Xu, G. Shi, Z. Liu, J. B. Rasmussen and J. D. Faris (2011). Two putatively homoeologous wheat genes mediate recognition of SnTox3 to confer effector-triggered susceptibility to *Stagonospora nodorum*. *The Plant Journal* **65,** 27–38.

Zhang, Z., E. Ersoz, C.-Q. Lai, R. J. Todhunter, H. K. Tiwari, M. A. Gore, P. J. Bradbury, J. Yu, D. K. Arnett, J. M. Ordovas et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics* **42,** 355–360.

Zhu, Z., F. Xu, Y. Zhang, Y. T. Cheng, M. Wiermer, X. Li and Y. Zhang (2010). Arabidopsis resistance protein SNC1 activates immune responses through association with a transcriptional corepressor. *Proceedings of the National Academy of Sciences* **107,** 13960–13965.

Appendices

A

Chapter 2 Supplementary Material

Table A.1 – *P. nodorum* effector scores for the AM panel

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
S1117_RAFFLES	S	GBR	1997	listed uk	4	0.5	4
S0117_MARIS_HALBERD	S	GBR	1972	genbank cz	2	3.25	4
S1591_AC_BARRIE	S	CAN	2003	listed uk	2.33	4	4
S0017_MARIS_ENSIGN	S	GBR	1968	genbank cz	0	4	4
S0185_SICCO	S	NLD	1973	genbank cz	4	4	4
S0248_HIGHBURY	S	GBR	1968	genbank cz	0	2	4
S0249_AINTREE	S	GBR	1981	genbank cz	1.5	2	4
S0337_SANDOWN	S	GBR	1980	genbank cz	2.33	4	0
S0406_NEWMARKET	S	GBR	NA	genbank cz	4	3.75	0
S0455_HAYDOCK	S	GBR	NA	genbank cz	3.5	3	0
S0494_SOLITAIRE	S	GBR	1985	genbank cz	1	4	1
S0497_JERICO	S	FRA	1983	listed uk	4	1	4
S0052_MARIS_DOVE	S	GBR	1971	genbank cz	4	2.5	4
S0053_MARIS_BUTLER	S	GBR	1972	genbank cz	1.25	0.25	4
S0547_WEMBLEY	S	GBR	1984	listed uk	3.25	0	0.25
S0919_CHABLIS	S	GBR	1994	listed uk	0	2	2
S0920_SHIRAZ	S	GBR	1994	listed uk	4	3	4
U100100_ANGLIA	W[1]	GBR	1987	genbank cz	0	3.5	1.5
U100136_ARINA	W[3]	CHE	1981	genbank cz	0	2.5	3.5
U100273_CAPELLE_DESPREZ	W	FRA	1946	genbank cz	NA	NA	NA
U100347_COURTOT	S[1]	FRA	1974	genbank cz	4	3	1
U100460_ETOILE_DE_CHOISY	W[1,2]	FRA	1950	genbank cz	1	4	2.5
U100546_GARNET	S[1]	CAN	1926	genbank cz	4	4	4
U100578_GREIF	W[1,2]	DEU	1989	genbank cz	0.5	3.75	2.25

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
U100825_MONOPOL	W[1,2,3]	DEU	1975	genbank cz	2	3	4
U100874_OKAPI	W[3]	DEU	1978	genbank cz	0	3	2
U101082_SELKIRK	S[1]	CAN	1953	genbank cz	NA	NA	NA
U101237_VIKING	W[1,2,3]	DNK	1956	genbank cz	2	4	3
U101246_VUKA	W[1,2,3]	DEU	1975	genbank cz	NA	NA	NA
U101415_IBIS	W[3]	DEU	1962	genbank cz	0	3.25	4
U112030_TAMBOR	W[1]	DEU	1995	genbank cz	0	4	1.75
U112495_ISENGRAIN	W[3]	FRA	1997	genbank cz	4	1.67	4
U112544_AMAROK	W[3]	FRA	NA	genbank cz	4	3.5	4
U112547_ARONDE	S[1]	FRA	1962	genbank cz	0	3.5	4
U112943_LORRAINE	W[1]	FRA	1998	genbank cz	0	0.67	0.25
U113096_BOSTON	W[1]	NLD	NA	NA	0	0.25	2.5
U9000_80_30_VERSAILLES	W[3]	FRA	NA	NA	4	3.75	3.5
U9001_BERSEE	W[1,2,3]	FRA	1936	genbank cz	1.13	4	1.13
U9002_BILBO	W[1]	GBR	1972	genbank cz	0.25	0	0.33
U9003_CARSTENS_V	W[3]	DEU	1921	genbank cz	NA	NA	NA
U9004_CARSTENS_VIII	W[1]	DEU	1952	genbank cz	0	0.75	0
U9005_CLEO	W[1,3]	NLD	1961	genbank cz	0	2	1.5
U9006_DW929620_511	NA	GBR	NA	NA	4	3.5	4
U9007_DW930861_509	NA	GBR	NA	NA	0	2	1
U9008_ELYSEE	S[1]	FRA	1968	genbank cz	0	0	1
U9010_HEINES_110	W[1,2]	DEU	1969	genbank cz	4	3.25	3.25
U9011_HEINES_PEKO	S[1]	DEU	1946	genbank cz	NA	NA	NA
U9012_HOLDFAST	W[1,2,3]	GBR	1936	genbank cz	1.5	0	3.99

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
U9013_HYBRID_46	W[1,3]	FRA	1936	genbank cz	0	2.25	1
U9014_HYBRIDE_DU_JONCQUOIS	W[1]	FRA	1933	genbank cz	0	4	0.75
U9015_JUFY_I	S[1]	BEL	1954	genbank cz	0.25	2.75	3.5
U9016_KOGA_1	S[1,3]	DEU	NA	NA	4	2	2
U9017_KOGA_II	S[1]	DEU	NA	NA	0	2	3.75
U9018_KRANICH	W[1,2]	DEU	1969	genbank cz	1	2	4
U9019_KRONJUWEL	W[1,2]	DEU	1981	genbank cz	1	0	0.5
U9020_MARNE_DESPREZ	W[1,2]	FRA	1954	genbank cz	NA	NA	NA
U9021_MINISTER	W[1,2]	BEL	1947	genbank cz	0.25	3.75	1.25
U9022_NORD_DESPREZ	W[1,3]	FRA	1945	genbank cz	0	0.75	0
U9023_NSL_92_5719	NA	GBR	NA	NA	0	0	3
U9024_NSL_93_5372	NA	GBR	NA	NA	0	0	1.25
U9025_NSL_94_5130	NA	GBR	NA	NA	0.33	0	1.33
U9026_NSL_94_6897	NA	GBR	NA	NA	0	0	0.25
U9027_NSL_WW13	NA	GBR	NA	NA	0.25	2.25	0
U9028_OBELISK	W[3]	NLD	1985	genbank cz	1.25	2.75	4
U9029_POROS	W[2]	DEU	1966	genbank cz	0.75	2	2
U9030_PROFESSEUR_DELOS	W[1,2]	BEL	1937	genbank cz	1	NA	NA
U9031_PROFESSEUR_MARCHAL	W[1,2,3]	BEL	1957	genbank cz	1.5	NA	NA
U9032_RIEBESEL_57/41	W[1]	DEU	NA	NA	4	1	4
U9033_STELLA	W[1]	BEL	1957	genbank cz	0	1	2.25
U9034_TADEPI	W[1,3]	FRA	1949	genbank cz	0	2.5	0
U9035_TADORNA	W[2,3]	NLD	1966	genbank cz	0	2.5	2
U9036_THATCHER	S[1,2]	USA	1934	genbank cz	NA	NA	NA

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
U9038_VILMORIN_23	W[1,2]	FRA	1923	genbank cz	NA	NA	NA
U9039_VILMORIN_27	W[3]	FRA	1928	genbank cz	NA	NA	NA
U9040_VILMORIN_29	W[1,2]	FRA	1929	genbank cz	0.5	2.25	NA
U9042_WH929623_5	NA	NA	NA	NA	0	3.5	1
U9043_YEOMAN	W[3]	FRA	1916	genbank cz	0	2.75	3.25
W0010_MILDRESS	W	NLD	1964	trial uk	1	3.33	3.5
W1003_VIVANT	W	GBR	1990	BR ref	4	4	4
W1008_JACADI	W	FRA	1997	listed uk	0.25	4	4
W1014_BANKER	W	GBR	1996	listed uk	0.25	4	1.25
W1019_TEMPLE	W	NA	1996	listed uk	0	3.75	0
W1020_TILBURI	W	FRA	1994	BR ref	4	3.5	2.75
W1028_TEMPEST	W	NA	1997	listed uk	0	2.5	2.75
W1029_BLAZE	W	GBR	1996	listed uk	0	2	4
W0103_MARIS_FREEMAN	W	GBR	1970	trial uk	0	2.25	0.25
W1030_CANTATA	W	GBR	1996	listed uk	0	1.5	0
W1031_FALSTAFF	W	GBR	1997	listed uk	0	4	4
W1033_SAVANNAH	W	GBR	1996	listed uk	0	1.67	1.5
W1035_MAVERICK	W	GBR	1996	listed uk	0	0.33	4
W1039_CHAUCER	W	GBR	1996	listed uk	0	0	0.67
W1040_WESTON	W	GBR	1996	listed uk	0	2.75	0.67
W1046_OBERON	W	GBR	1996	listed uk	0	2.75	2
W1047_KRAKATOA	W	GBR	1996	listed uk	0	2.5	3.75
W1048_WELLINGTON	W	GBR	1997	listed uk	1.5	1.75	0
W0105_MARIS_MARKSMAN	W	GBR	1975	trial uk	0.25	1.25	2.33

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W1062_SHANGO	W	GBR	1995	trial uk	0	2.75	1
W1070_CLAIRE	W	GBR	1997	listed uk	0.5	2.15	1.01
W1078_MARSHAL	W	GBR	1997	listed uk	0	3	3.25
W1088_ROSETTE	W	NA	1997	listed uk	0.25	0.75	4
W0109_MEGA	W	GBR	1970	trial uk	0	4	0
W1092_SHAMROCK	W	FRA	1997	listed uk	0	4	4
W0110_CARIBO	W	DEU	1968	genbank cz	0.75	0	1.25
W1100_BUCHAN	W	GBR	1997	listed uk	NA	NA	NA
W1108_DATUM	W	GBR	1999	listed uk	0	2	3
W0111_ATOU	W	FRA	1970	trial uk	0	3.5	4
W1111_AARDVARK	W	GBR	1997	listed uk	0	0	1.75
W1115_DICKINS	W	NA	1998	listed uk	4	0	2
W1120_FLAIR	W	DEU	1996	BR ref	0	0	0
W1122_ATOLL*	W	FRA	1996	BR ref	0	0	1
W1126_ECLIPSE	W	GBR	1999	listed uk	0	0	4
W1133_AGAMI	W	GBR	1996	BR ref	0	0	2.75
W1137_BROILER	W	GBR	1999	listed uk	0	0.25	0
W1147_NAPIER	W	GBR	1998	listed uk	0	0	1
W1149_WICKHAM	W	GBR	1998	listed uk	0	0	0
W1151_CHICAGO	W	GBR	1998	listed uk	0	0	3.5
W1152_EXPLOSIV	W	GBR	1999	listed uk	0	0	0
W1153_GENGHIS	W	GBR	1999	listed uk	0.5	3.5	4
W1155_DENVER*	W	GBR	1998	listed uk	0	2	2.5
W1156_BRANDO	W	GBR	1999	listed uk	0	4	4

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W1157_RANGER	W	GBR	1998	listed uk	0.75	3.5	3.33
W1161_DERWENT*	W	USA	1999	listed uk	0	3	4
W1169_MEXICO	W	NA	1998	listed uk	0	2	2.75
W1184_ARLINGTON	W	USA	2000	listed uk	0	2.75	3.5
W1195_VERDON	W	FRA	1999	listed uk	1	2	2.33
W1196_ELECTRON	W	FRA	2000	listed uk	0	0.67	4
W1197_DORIAL	W	FRA	1999	listed uk	0	0	0
W1199_MILESTONE	W	GBR	1999	listed uk	NA	NA	NA
W1200_OPTION	W	GBR	1999	listed uk	0	1.5	2
W1201_REYDON	W	GBR	1999	listed uk	0	2.25	2
W1202_OXBOW	W	GBR	1999	listed uk	0	1.25	1
W1204_SLADE	W	GBR	1999	listed uk	0	2.5	2.33
W1205_POTENT	W	GBR	1999	listed uk	0	2	3.75
W1206_GOLDLACE	W	GBR	1999	listed uk	0.5	2.5	4
W1210_ORTON	W	GBR	2000	listed uk	0	0.5	0
W1211_ODYSSEY	W	GBR	1999	listed uk	0	1.75	3.5
W1213_EXSEPT	W	DEU	1999	listed uk	1.25	3.5	4
W1220_DEBEN	W	GBR	1999	listed uk	0	2.25	2
W1222_ARRIVA	W	GBR	1999	listed uk	0.25	4	1.75
W1223_CANTERBURY	W	GBR	1999	listed uk	0	0.5	0
W1224_CYBER	W	GBR	1997	BR ref	4	2.5	4
W1226_TRAVIX	W	GBR	2000	listed uk	0	2.5	4
W1227_BISCAY	W	GBR	1999	listed uk	0	0	1.67
W1233_ARK	W	DEU	2000	listed uk	0	0	0.67

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W1235_EPOCH	W	NA	1997	BR ref	0	1.5	4
W1240_NEXUS	W	NA	1997	BR ref	0	0	0
W1249_CHEQUER	W	GBR	1998	BR ref	0	1.25	0.67
W1251_HARROW	W	GBR	1998	BR ref	0.25	1.5	2.5
W1257_RICHMOND	W	GBR	2000	listed uk	0.33	1.5	1.75
W1258_CHATSWORTH	W	NLD	2000	listed uk	0	0	0.33
W1259_WOBURN	W	NLD	2000	listed uk	1	3.25	4
W1262_MACRO	W	GBR	2000	listed uk	NA	2	2.5
W1263_FENDER	W	GBR	2001	listed uk	1.5	2.75	2.75
W1266_ACCESS	W	GBR	2000	listed uk	NA	2.75	3.5
W1267_ANGLO	W	GBR	2000	listed uk	NA	2.75	2.5
W1269_POSIT	W	GBR	2000	listed uk	2	4	4
W1270_COMET	W	GBR	1998	BR ref	2	3.75	4
W1271_FEAST	W	GBR	2000	listed uk	1.75	4	4
W1272_RAMPART	W	GBR	2000	listed uk	1.33	2.5	4
W1273_VAULT	W	GBR	1998	BR ref	1	2	2.67
W1275_KEMPT	W	GBR	1998	BR ref	0	4	4
W1276_ALCHEMIST	W	GBR	2000	listed uk	NA	4	4
W1277_FRELON	W	GBR	2001	listed uk	4	2	0
W1278_PHLEBAS	W	GBR	2000	listed uk	0	3.75	0
W1280_A13_98	W	DNK	NA	NA	0	2	4
W1281_XI19	W	NLD	2000	listed uk	4	3.87	4
W1282_SOLSTICE*	W	NLD	2001	listed uk	3.75	1.75	4
W1286_STORM	W	NLD	2000	listed uk	0	2	4

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W1292_VIRTUOSE	W	FRA	1998	BR ref	0	1	4
W1295_SABRE*	W	GBR	2000	listed uk	0	0	0.5
W1300_WINDSOR	W	DEU	1998	BR ref	0	2.25	4
W1310_CAPNOR	W	FRA	1999	BR ref	0	0.5	4
W1321_PR21R60	W	NA	2001	listed uk	0	1	4
W1326_TELLUS	W	GBR	2002	listed uk	0	2.75	4
W1327_BRUNEL	W	GBR	2001	listed uk	0	2	4
W1328_WIZARD***	W	GBR	2002	listed uk	0	0	0.5
W1329_ZAKA	W	GBR	2002	listed uk	0	2	2.75
W0133_TALENT	W	FRA	1973	genbank cz	0	0	0
W1330_ROBIGUS	W	GBR	2002	listed uk	0	1.03	1.5
W1331_CONTEXT	W	GBR	2001	listed uk	0	2.33	2
W1332_CONVOY	W	GBR	2001	listed uk	0	4	4
W1333_ASHANTI	W	GBR	2001	listed uk	0	1.67	0
W1335_SCORPION_25	W	GBR	2001	listed uk	4	4	4
W1336_WARLOCK_24	W	GBR	2001	listed uk	4	4	4
W0134_CLEMENT	W	NLD	1971	trial uk	0	1	2
W1343_CARLTON	W	GBR	2001	listed uk	4	0.67	3
W1353_GOODWILL	W	GBR	2001	listed uk	0	2.5	3
W1359_INSIGHT	W	GBR	2001	listed uk	0	2	1.75
W1361_MALLET	W	GBR	2001	listed uk	0	0.5	3
W1364_RIVET	W	GBR	2001	listed uk	0	1.5	0.5
W1367_FIELDER	W	GBR	2001	listed uk	0	1	4
W1370_HARBOUR	W	GBR	2001	listed uk	0	1.67	NA

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W1376_EINSTEIN	W	GBR	2002	listed uk	0	0.75	2
W1380_ARRAN	W	GBR	2002	listed uk	0	1.33	1
W1383_CHARDONNAY	W	GBR	2000	trial uk	0	0.25	2.67
W1388_CORDIALE	W	GBR	2003	listed uk	0	3.63	0.5
W1392_QUEST	W	GBR	2003	listed uk	0.25	1.25	0.25
W1395_SENATOR	W	GBR	2003	listed uk	0.67	2.67	2.25
W1400_FLAXEN	W	GBR	2002	listed uk	0	1.5	0.67
W1407_AWARD	W	DEU	2002	listed uk	0.25	0.67	2
W1409_MAYFAIR	W	NA	2002	listed uk	0	2.25	NA
W1411_DICKSON	W	NA	2002	listed uk	0	2.25	1.25
W1414_VECTOR	W	GBR	2002	listed uk	0	0.33	3.33
W1415_SMUGGLER	W	GBR	2002	listed uk	0	2.67	NA
W1426_ISTABRAQ	W	GBR	2003	listed uk	0	2.67	1.5
W1427_NIJINSKY	W	GBR	2003	listed uk	0	2.75	0.25
W1428_SANCERRE	W	NA	2003	listed uk	0	1.75	0
W1429_ALSACE	W	FRA	2002	listed uk	0.25	2	2.75
W1435_SW_TATAROS	W	DEU	2003	listed uk	1.5	4	1
W1438_CAPHORN	W	GBR	2000	BR ref	0	4	4
W1439_DART*	W	GBR	2002	listed uk	0	1.25	1.33
W1441_MONUMENT	W	GBR	2002	listed uk	0	2	2
W1442_GLADIATOR*	W	GBR	2002	listed uk	0	2.28	4
W1443_HERITAGE	W	GBR	2002	listed uk	0	2	3.25
W1444_M007	W	NA	2002	listed uk	0	3.25	4
W1445_BELTER	W	GBR	2002	listed uk	0.25	2	2

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W1447_BENTLEY	W	FRA	2002	listed uk	NA	NA	NA
W1449_WELFORD	W	GBR	2002	listed uk	0	1.67	3.5
W1450_PENNANT*	W	GBR	2002	listed uk	0	0	2
W1456_RAGLAN	W	GBR	2003	listed uk	0	1.33	3.25
W1461_MAYFIELD	W	GBR	2003	listed uk	0	1.75	4
W1462_AMBROSIA	W	GBR	2003	listed uk	0	0	0
W1463_EQUATOR	W	GBR	2003	listed uk	0	2	3
W1464_EXTEND	W	GBR	2003	listed uk	0	1.75	1
W1467_MAGNITUDE	W	GBR	2003	listed uk	0	1.25	3.25
W1468_DEFENDER	W	GBR	2003	listed uk	0	2.5	3
W1469_SCANDIA	W	GBR	2003	listed uk	0	4	3.25
W1470_CHOICE*	W	GBR	2003	listed uk	0.25	4	2
W1477_ATLANTA	W	GBR	2004	listed uk	0	0	0.5
W1482_GLASGOW	W	NA	2003	listed uk	0	2	4
W1489_ISIDOR	W	FRA	2002	listed FRA	0	0.5	0
W1497_SAMURAI	W	DEU	2003	listed uk	0	2.25	4
W1499_PREDATOR	W	GBR	2003	listed uk	0	3.33	3.75
W1502_BROMPTON	W	GBR	2002	trial uk	0	1.12	2.62
W1511_CHESTER	W	GBR	2003	listed uk	0	0	0.25
W1512_EXETER	W	FRA	2003	listed uk	0.5	1.5	0.5
W1532_MASCOT*	W	GBR	2004	listed uk	0.75	2.5	1.67
W1541_PIRANHA	W	GBR	2004	listed uk	0.25	1.25	4
W1542_DIRECTOR	W	GBR	2004	listed uk	1	1	2.67
W1543_ASAGAI	W	GBR	2003	trial uk	0.5	1.25	4

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W1545_ZEBEDEE	W	GBR	2004	listed uk	1	1.33	2.67
W1546_GATSBY	W	GBR	2004	listed uk	0	2	4
W1549_FASTNET	W	GBR	2004	listed uk	0.5	4	2.5
W1550_DOVER	W	GBR	2005	listed uk	0	0	0.33
W1556_HOURRA	W	GBR	2004	listed uk	1.25	2	2
W1561_HYPERION	W	FRA	2004	listed uk	0.5	2.33	0
W1564_ALCHEMY	W	GBR	2004	listed uk	0	2.12	0.37
W1573_AARDEN	W	FRA	2005	listed uk	0	0	0
W1577_HURLEY	W	GBR	2005	listed uk	0	1.25	0
W1578_KIPLING	W	FRA	2004	listed uk	0	1.75	0
W0158_MARIS_FUNDIN	W	GBR	1972	trial uk	0	3.25	0
W1599_BATTALION	W	GBR	2005	listed uk	0	4	1
W1606_OCHRE	W	GBR	2003	BR ref	0	0.5	0
W1611_CONTENDER	W	GBR	2006	listed uk	0	0	4
W1621_GULLIVER	W	GBR	2005	listed uk	0	2	4
W1625_BENEDICT	W	GBR	2006	listed uk	0	3	1.67
W1649_SAHARA	W	NLD	2005	listed uk	0	4	1.75
W1652_HUMBER	W	GBR	2005	listed uk	0	4	1
W1658_OAKLEY	W	GBR	2006	listed uk	1.5	2.25	4
W1666_CANADAIR	W	NA	2006	listed uk	1.33	4	0
W1667_BUZZER	W	GBR	2006	listed uk	0.33	4	4
W1668_MAXWELL	W	NA	2006	listed uk	0	0.67	0
W1677_ROCKY*	W	GBR	2005	trial uk	0	3.25	2.75
W1680_MUSKETEER	W	GBR	2005	trial uk	0	3.75	0.5

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W1687_MARKSMAN**	W	GBR	2006	listed uk	0	4	2.5
W0169_ALCEDO	W	DEU	1974	genbank cz	4	3.25	0.5
W1695_BOWINDO	W	GBR	2007	listed uk	0	4	2
W1704_VELOCITY	W	GBR	2006	listed uk	3	2.5	4
W0171_HARDI	W	NA	1969	genbank cz	0.67	3	0
W1714_ORATOR	W	FRA	2007	listed uk	0	3	4
W1724_EMERALD	W	GBR	2007	listed uk	1	4	0.67
W1725_DUXFORD	W	GBR	2006	listed uk	3.25	4	4
W1726_LIMERICK	W	GBR	2006	listed uk	4	3.25	4
W1727_MONTY	W	GBR	2006	listed uk	0	3.25	2
W1728_ZANATAN	W	GBR	2006	listed uk	0.25	4	1.5
W0173_FLANDERS	W	FRA	1973	trial uk	1.5	NA	4
W1731_HEREFORD	W	DNK	2005	trial uk	0	4	0.33
W1737_JB_DIEGO	W	GBR	2006	listed uk	0	3	1
W1746_WALPOLE	W	NLD	2005	BR ref	4	2	4
W1753_BATSMAN	W	NA	2005	BR ref	0	2	4
W1756_LANGDALE	W	GBR	2005	BR ref	0	1.75	4
W1760_GALTIC	W	FRA	2005	BR ref	0	4	4
W1766_GALLANT	W	GBR	2005	BR ref	0	2.75	0.75
W1769_SHOGUN	W	GBR	2005	BR ref	0	0	2
W1779_CELEBRATION	W	NA	2005	BR ref	0	0.75	1.25
W0178_KINSMAN	W	GBR	1973	trial uk	0	3.75	4
W1787_SCOUT	W	GBR	2005	BR ref	1	0	1
W1789_QPLUS	W	NA	2005	BR ref	0.25	0	4

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W0179_HOBBIT	W	GBR	1973	trial uk	1.5	0	2.33
W1790_LEAR	W	GBR	2005	BR ref	1.25	0.75	3.5
W1795_BANTAM*	W	GBR	2005	BR ref	4	2	4
W1798_CASSIUS	W	GBR	2005	BR ref	0.25	0.25	1.25
W0018_MARIS_RANGER	W	GBR	1966	trial uk	2	4	2.75
W1801_PANORAMA	W	GBR	2005	BR ref	4	2	4
W1806_TIMARU	W	GBR	2005	BR ref	0	0	4
W1808_ACROBAT	W	GBR	2005	BR ref	0	0	1.75
W1811_GRAFTON	W	GBR	2005	BR ref	0	3.5	0.25
W1813_CONQUEROR	W	GBR	2005	BR ref	1.5	0	2
W1825_TYRELL	W	GBR	2006	BR ref	NA	NA	NA
W1827_CHEVRON**	W	GBR	2006	BR ref	0	3.25	4
W1830_ROCHFORT	W	GBR	2006	BR ref	0	0	3.5
W1853_INVICTA	W	GBR	2006	BR ref	0	0	3.25
W1858_SANTANA	W	GBR	2006	BR ref	0	2	4
W1860_EDMUNDS	W	GBR	2006	BR ref	0	0	0
W1865_WARRIOR**	W	GBR	2006	BR ref	0	0	2.67
W1871_STALWART	W	GBR	2006	BR ref	0.5	0	1.5
W1877_KWS_CURLEW	W	GBR	2006	BR ref	4	2	4
W1880_KWS_STERLING	W	GBR	2006	BR ref	0	1	1.25
W1882_HORIZON	W	GBR	2006	BR ref	0	0.25	1.25
W1883_KWS_QUARTZ	W	GBR	2006	BR ref	NA	NA	NA
W1885_BELUGA	W	GBR	2006	BR ref	0	0.67	0.75
W1892_CADOGAN	W	GBR	2007	BR ref	0	0.25	2.25

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W1895_DENMAN	W	GBR	2007	BR ref	0	0	0
W1904_RAINBOW*	W	DEU	2007	BR ref	0.5	1.25	4
W1907_KWS_PODIUM	W	GBR	2007	BR ref	0.5	3	0
W1909_KWS_GYMNAST	W	GBR	2007	BR ref	0	0.5	1.25
W1911_KWS_TARGET	W	GBR	2007	BR ref	0	2.75	0
W1916_KWS_SANTIAGO	W	GBR	2007	BR ref	1.5	4	4
W0192_DURIN	W	GBR	1968	genbank cz	2.5	0.75	3
W1922_COCOON	W	GBR	2007	BR ref	0.5	1.75	4
W0193_TIPSTAFF	W	GBR	NA	NA	0	4	2.75
W1933_ORBIT	W	GBR	2007	BR ref	2	4	1.5
W1940_GRAVITAS	W	GBR	2007	BR ref	0.5	4	2
W1941_STIGG	W	GBR	2007	BR ref	0	1.5	0
W1943_LAZARUS	W	GBR	2007	BR ref	0	4	0
W1947_SHELDON	W	GBR	2007	BR ref	2	4	2
W1954_TUXEDO	W	GBR	2007	BR ref	1	2.25	4
W0196_SPORTSMAN	W	GBR	1974	trial uk	0	4	2
W0201_ARMADA	W	GBR	1974	trial uk	0	2.75	4
W0205_KADOR	W	GBR	1974	trial uk	0.75	2.5	3
W0023_TOMMY	W	FRA	1967	trial uk	0.5	2	1.25
W0230_HUSTLER	W	GBR	1975	trial uk	0	0	0
W0231_BRIGAND	W	GBR	1975	trial uk	3.25	3	4
W0233_MARDLER	W	GBR	1975	trial uk	0	4	3.33
W0235_ARGENT	W	GBR	1979	genbank cz	0.5	1	3
W0243_AQUILA	W	GBR	1975	trial uk	NA	NA	NA

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W0260_ANVIL	W	GBR	1982	genbank cz	0	0	0
W0265_COPAIN	W	FRA	1976	trial uk	2	0.5	1.68
W0270_WIZARD	W	GBR	1983	genbank cz	0	0	1.75
W0271_IONA	W	GBR	1976	trial uk	1.25	0.25	2
W0272_SENTRY	W	GBR	1976	trial uk	0.25	0	1
W0273_VILLEIN	W	DEU	NA	NA	1	0	2.5
W0274_BOUNTY	W	GBR	1976	trial uk	0.75	0	3
W0276_VIRTUE	W	GBR	1976	trial uk	1	0	1.5
W0286_HERALD	W	GBR	NA	NA	0.25	0	1.75
W0287_AVALON	W	GBR	1979	trial uk	0	0.5	0.5
W0288_MAESTRO	W	GBR	1976	genbank cz	0	1.33	2
W0289_SHIRE	W	GBR	1979	genbank cz	0	1.25	1
W0291_HEDGEHOG	W	GBR	NA	NA	0.25	1.5	2.75
W0296_GRANTA	W	GBR	1980	genbank cz	0	3.33	0.67
W0311_ABELE	W	GBR	1978	trial uk	0	0.5	0
W0317_CHAMPLEIN	W	FRA	1959	trial uk	0	1	0.67
W0319_MARIS_WIDGEON	W	GBR	1960	trial uk	0	0.67	1.67
W0321_NORMAN	W	GBR	1979	trial uk	0.5	2.5	0.67
W0323_FLAMBEAU	W	GBR	NA	NA	0	2	0.33
W0325_RAPIER	W	GBR	1978	trial uk	0	3	3.25
W0328_DISPONENT	W	DEU	1975	genbank cz	1	0.5	3
W0364_LONGBOW	W	GBR	1979	trial uk	1	2	0
W0371_JENA	W	NA	1980	genbank cz	1.33	3.25	4
W0389_STETSON	W	GBR	1980	trial uk	0	2	NA

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W0039_MARIS_SETTLER	W	GBR	1967	trial uk	0	3.25	0
W0004_BOUQUET	W	FRA	1965	trial uk	4	2.75	2.75
W0041_MARIS_BEACON	W	GBR	1967	trial uk	0.67	3.33	4
W0042_MARIS_NIMROD	W	GBR	1968	trial uk	0.5	0.67	4
W0423_HAMMER	W	GBR	1981	trial uk	1.33	1.33	0
W0043_MARIS_ENVOY	W	GBR	1974	genbank cz	0	NA	1
W0438_MITHRAS	W	GBR	1978	trial uk	2	2	0
W0439_AVOCET	W	GBR	NA	NA	1.75	4	0
W0440_GALAHAD	W	GBR	1982	listed uk	0.88	2.25	0
W0464_FRONTIER	W	GBR	1981	genbank cz	4	1.75	3.75
W0479_AMBASSADOR	W	GBR	1982	trial uk	0.25	3.5	4
W0486_MOULIN	W	GBR	1980	trial uk	0.5	2.5	3.5
W0488_RENARD	W	GBR	1982	trial uk	1	4	3.67
W0489_BROCK	W	GBR	1982	trial uk	1.67	2.25	3.5
W0491_BRIMSTONE	W	GBR	1982	trial uk	1	3.67	2.5
W0496_TONIC	W	GBR	1983	listed uk	4	1	2.5
W0509_BOXER	W	GBR	1982	trial uk	0	1.67	1
W0527_GAWAIN	W	GBR	1984	trial uk	4	2.75	4
W0528_ASLAN	W	GBR	1982	genbank cz	3.5	1.25	4
W0529_CRAFTSMAN	W	GBR	1987	genbank cz	1	2.25	1
W0531_BOOTY	W	GBR	NA	NA	0.75	1	4
W0533_MERCIA	W	GBR	1984	listed uk	0.25	0	0
W0537_SLEJPNER	W	SWE	1983	trial uk	0.82	1	0
W0544_AXONA	W	NLD	1984	listed uk	4	3.75	4

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W0565_PARADE	W	GBR	1984	trial uk	0	1	1
W0584_SARSEN	W	GBR	1987	genbank cz	0.5	1.33	1
W0585_RENDEZVOUS	W	GBR	1984	trial uk	1	1.67	1.75
W0590_CINNABAR	W	GBR	1985	genbank cz	0	1.5	0.75
W0591_HORNET	W	GBR	1984	trial uk	0	1.67	0.33
W0607_APOLLO	W	DEU	1985	trial uk	0	0.75	0.67
W0610_SOLEIL	W	FRA	1986	listed uk	0	0	NA
W0612_ARMINDA	W	NLD	1976	genbank cz	2.5	2.67	0
W0618_MERLIN	W	DEU	1956	trial uk	0	4	3
W0626_SQUADRON	W	GBR	NA	NA	0.75	2.75	0
W0628_RIBAND	W	GBR	1987	listed uk	0	0.67	0
W0063_BENNO	W	DEU	1969	trial uk	0	0	4
W0635_MANDATE	W	GBR	1985	trial uk	1.25	2.25	0
W0065_MARIS_PLOUGHMAN	W	GBR	1972	genbank cz	0.25	1.75	1.25
W0066_MARIS_HUNTSMAN	W	GBR	1969	trial uk	0	1	0.33
W0067_MARIS_TEMPLAR	W	GBR	1969	trial uk	0	1.75	4
W0670_APOSTLE	W	GBR	1986	trial uk	1.5	1.25	3
W0671_PASTICHE	W	GBR	1988	listed uk	0.5	3	0
W0672_FRESCO	W	GBR	1988	listed uk	1.25	2.25	4
W0682_URBAN	W	DEU	1988	trial uk	1.67	1	4
W0692_BEAVER	W	GBR	1989	listed uk	0	3	2
W0694_HAVEN	W	GBR	1988	listed uk	0.25	3.25	4
W0707_AXIAL	W	FRA	1988	trial uk	2.5	2.5	3
W0724_EKLA	W	FRA	1988	genbank cz	1.5	3.5	3.75

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W0728_TALON	W	DEU	1988	trial uk	0.5	2.5	4
W0732_DEAN	W	GBR	1987	trial uk	0	2.75	4
W0735_TARA	W	GBR	1988	trial uk	0	3	3.25
W0736_HEREWARD	W	GBR	1989	listed uk	1	2.37	4
W0755_ORESTIS	W	DEU	1988	genbank cz	0	1.5	4
W0756_SITKA	W	GBR	NA	NA	0	1.25	0
W0759_ADMIRAL	W	GBR	1990	listed uk	0	0	2
W0760_SOLDIER	W	GBR	NA	NA	0.25	1.5	2
W0765_SAREK	W	SWE	1992	genbank cz	4	2	4
W0770_LANCELOT	W	NA	2002	listed FRA	0.67	2	3.5
W0773_TORONTO	W	DEU	1990	listed DEU	1	0	0.75
W0775_ESTICA	W	NLD	1990	listed uk	0	2	3
W0776_WASP	W	GBR	1990	listed uk	2.75	1	4
W0783_CAPRIMUS	W	GBR	1994	genbank cz	0	0	4
W0785_ARISTOCRAT	W	GBR	1992	genbank cz	NA	2	4
W0787_TORFRIDA	W	GBR	1988	trial uk	1	0.5	4
W0798_CAMP_REMY	W	FRA	1980	listed FRA	0	0	1
W0801_ASTRON	W	DEU	1990	BR ref	0	4	4
W0803_FLETUM	W	NLD	1990	BR ref	0	2.5	4
W0806_DIABLO	W	GBR	1990	BR ref	0.75	2	1.33
W0808_SPARK	W	GBR	1991	listed uk	4	2	3.5
W0810_ZODIAC	W	GBR	1990	trial uk	0	1.75	0.67
W0811_FENDA	W	GBR	1990	BR ref	0.67	4	4
W0817_HUSSAR	W	GBR	1991	listed uk	0.5	1.5	1.25

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W0818_BRIGADIER	W	GBR	1991	listed uk	1.38	0.63	2.71
W0824_BRYDEN	W	GBR	1990	BR ref	0.25	2.5	0.5
W0825_ADROIT	W	GBR	1990	BR ref	0.75	2.75	4
W0826_LEO*	W	NA	1990	BR ref	0	0.33	0.5
W0827_RENOWN	W	NA	1990	BR ref	0.5	4	4
W0828_HUNTER	W	GBR	1991	listed uk	0	1.67	2.67
W0829_VERITAS	W	GBR	1990	BR ref	0	4	4
W0833_CADENZA	W	GBR	1991	listed uk	4	4	4
W0834_SOISSONS	W	FRA	1990	listed uk	3.72	2.87	4
W0844_TREND	W	DEU	1990	BR ref	0	4	3
W0847_FLAME	W	GBR	1992	listed uk	0	3	0.5
W0848_WOODSTOCK	W	GBR	1992	listed uk	0	4	2
W0856_LYNX*	W	GBR	1992	listed uk	0	3.5	1.25
W0858_RIALTO	W	GBR	1992	listed uk	0.25	3.19	4
W0859_ANDANTE	W	GBR	1991	trial uk	0.5	1	1.33
W0862_ATLA	W	GBR	1990	BR ref	0	2	4
W0863_OSTARA	W	NA	1990	BR ref	0	0	0
W0867_PROPHET	W	GBR	1991	trial uk	0	2.5	0
W0878_BEAUFORT	W	GBR	1993	listed uk	0	2	0.75
W0881_ENCORE	W	GBR	1993	listed uk	4	2.67	2
W0882_CONSORT	W	GBR	1993	listed uk	0	2	2
W0886_TURPIN	W	GBR	1991	BR ref	0	2	0
W0896_DYNAMO	W	GBR	1993	listed uk	0	2	1
W0900_TRAWLER	W	GBR	1991	BR ref	0	1	3

Continued on next page

Table A.1 – continued from previous page

Ordered Variety	SGH	Origin	Year	Year Source)	SnToxA	SnTox1	SnTox3
W0931_CAXTON	W	GBR	1994	listed uk	0	2	3.75
W0932_REAPER	W	GBR	1994	listed uk	0	2	0.33
W0933_RALEIGH	W	GBR	1995	listed uk	0	3	1.75
W0934_DRAKE	W	GBR	1996	listed uk	0	2	4
W0938_SHANNON	W	GBR	1992	BR ref	0	0.75	3.75
W0939_CHARGER	W	GBR	1994	listed uk	0	2.75	3.33
W0940_CLOVE	W	GBR	1995	listed uk	0	2.25	3.75
W0942_NEWHAVEN	W	NA	1992	BR ref	0	1	2
W0943_GALATEA	W	GBR	1995	listed uk	0	1	0.75
W0945_MAGELLAN	W	GBR	1995	listed uk	0.33	1.5	3.75
W0946_CHIANTI	W	GBR	1995	listed uk	0	1.33	1
W0954_CROFTER	W	GBR	1994	listed uk	0	0.67	0
W0955_RITMO	W	NLD	1992	BR ref	0	1.75	4
W0966_RUBENS	W	FRA	1993	BR ref	2	0	0
W0967_ABBOT	W	GBR	1995	listed uk	1.33	1.75	2.75
W0970_IMPALA	W	GBR	1996	listed uk	0.5	1	0.5
W0972_ACCLAIM	W	GBR	1995	listed uk	0	2	2
W0973_MADRIGAL	W	GBR	1995	listed uk	0.5	0.67	4
W0978_HARRIER	W	GBR	1995	listed uk	0	2.5	0.5
W0980_MALACCA	W	GBR	1995	listed uk	0.5	4	0
W0982_HUDSON	W	GBR	1993	BR ref	0	4	0
W0983_EQUINOX	W	GBR	1995	listed uk	1.33	4	2
W0992_HOLSTER	W	GBR	1993	BR ref	0	2.5	4
W0994_WARRIOR	W	GBR	1995	listed uk	0	1	2

Association mapping panel SnToxA, SnTox1 and SnTox3 phenotypic data (mean of 3-4 replicates per variety). Score 0 = insensitive, score 4 = highly sensitive. SGH = seasonal growth habit (S=spring, W=winter. [1]Our phenotype data. [2]Source = European Wheat Database, http://genbank.vurv.cz/ewdb/. [3]Other online sources). NA = missing data. Year = consensus varietal release date, based on the information listed in column 'Year source'. BR control = brown rust control variety.

† Listed UK = first year on UK National List; trial UK = first year in trial for UK Recommended List; genebank cz = year listed in the European Wheat Database, http://genbank.vurv.cz/ewdb/.

B

Chapter 3 Supplementary Material

Table B.1 – Protein domains of the predicted genes in the Cadenza *Snn3-B1* pseudomolecule

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g1					
g2					
g3					
g4	superfamily	3.94E-05	44.2813	cl04444	Transposase_28 superfamily
g5					
g6					
g7	superfamily	5.74E-19	81.3648	cl14782	RNase_H_like superfamily
	superfamily	2.56E-08	51.111	cl21549	rve superfamily
g8					
g9					
g10	specific	2.77E-12	58.8223	cd00018	AP2
g11	superfamily	7.82E-05	42.1803	cl10012	DnaQ_like_exo superfamily
g12	superfamily	5.54E-09	51.9112	cl26047	Retrotran_gag_2 superfamily
g13					
g14	superfamily	9.02E-06	44.2476	cl29674	Retrotrans_gag superfamily
g15	superfamily	1.49E-03	35.4928	cl34947	COG5222 superfamily
g16					
g17	specific	1.46E-27	103.152	pfam12274	DUF3615
g18	superfamily	0	553.311	cl27418	Branch superfamily
g19	superfamily	1.43E-06	44.8851	cl29659	AAA superfamily
	superfamily	3.70E-04	39.4303	cl32052	PRK03992 superfamily
g20	specific	1.78E-27	106.088	pfam14363	AAA_assoc
	specific	1.07E-106	329.597	cd17354	MFS_Mch1p_like
g21	specific	5.06E-18	79.553	pfam00004	AAA

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
	specific	4.84E-29	108.785	pfam14363	AAA_assoc
g22	specific	7.78E-18	80.3234	pfam00004	AAA
	superfamily	2.71E-08	52.9742	cl29659	AAA superfamily
	specific	2.28E-28	109.17	pfam14363	AAA_assoc
	superfamily	7.40E-05	46.0672	cl35251	PRK04195 superfamily
g23	specific	2.86E-20	84.6936	pfam03732	Retrotrans_gag
g24	superfamily	4.09E-04	42.6019	cl16623	DUF4283 superfamily
	superfamily	2.01E-11	66.2209	cl00490	EEP superfamily
	superfamily	9.95E-09	61.1077	cl33720	PHA03247 superfamily
	superfamily	9.87E-05	44.4125	cl31762	PTZ00368 superfamily
	specific	2.21E-12	65.7486	pfam00665	rve
	specific	1.15E-16	76.5371	pfam13966	zf-RVT
g25					
g26	superfamily	9.65E-50	163.673	cl29371	Transposase_21 superfamily
g27					
g28					
g29					
g30					
g31					
g32	specific	3.03E-19	80.8012	pfam14223	Retrotran_gag_2
g33	superfamily	2.68E-09	54.0113	cl14782	RNase_H_like superfamily
	superfamily	3.99E-25	99.3659	cl06662	RVT_2 superfamily
g34	superfamily	2.67E-04	39.2737	cl04444	Transposase_28 superfamily
g35					

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g36	superfamily	1.41E-07	53.1367	cl37069	SMC_prok_B superfamily
g37	specific	1.65E-12	63.5076	pfam03732	Retrotrans_gag
g38	superfamily	4.11E-05	44.6623	cl37069	SMC_prok_B superfamily
g39	superfamily	1.63E-05	41.168	cl11403	pepsin_retropepsin_like superfamily
g40					
g41					
g42	superfamily	8.51E-03	37.8151	cl03844	Folate_rec superfamily
	superfamily	7.40E-27	112.254	cl26817	GSDH superfamily
g43					
g44					
g45	superfamily	1.69E-03	40.6812	cl28772	Peptidase_S46 superfamily
g46	superfamily	5.75E-04	39.6272	cl11403	pepsin_retropepsin_like superfamily
	specific	1.13E-13	67.3596	pfam03732	Retrotrans_gag
g47	superfamily	3.94E-04	43.5179	cl37070	SMC_prok_A superfamily
g48	superfamily	7.49E-04	39.6272	cl11403	pepsin_retropepsin_like superfamily
g48	specific	1.07E-36	134.522	cd09279	RNase_HI_like
	specific	1.33E-52	182.025	cd01647	RT_LTR
g49	superfamily	2.40E-57	176.327	cl21562	DDE_Tnp_4 superfamily
g50	superfamily	3.35E-03	38.2872	cl34772	CobT2 superfamily
g51	superfamily	2.79E-04	43.1215	cl37069	SMC_prok_B superfamily
g52	superfamily	4.72E-08	55.0739	cl37070	SMC_prok_A superfamily
	superfamily	2.16E-03	38.5033	cl04444	Transposase_28 superfamily
g53					
g54					

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g55					
g56					
g57	superfamily	4.77E-20	78.2832	cl14782	RNase_H_like superfamily
g58					
g59					
g60	superfamily	0.000344522	43.5496	cl34174	Smc superfamily
g61					
g62					
g63					
g64	specific	9.31E-17	75.7936	pfam14223	Retrotran_gag_2
	superfamily	3.45E-04	39.2748	cl06662	RVT_2 superfamily
g65					
g66	superfamily	1.55E-04	44.7822	cl37532	DUF390 superfamily
g67	specific	4.21E-07	47.3312	cd00303	retropepsin_like
g68	superfamily	5.08E-08	56.5865	cl35953	rne superfamily
g69	superfamily	9.17E-05	45.4756	cl34174	Smc superfamily
g70	superfamily	1.87E-03	40.1129	cl25815	Spc7_N superfamily
g71					
g72					
g73	superfamily	4.95E-07	54.1115	cl36108	PRK12323 superfamily
	superfamily	3.04E-06	46.1518	cl16504	Transpos_assoc superfamily
	superfamily	1.42E-80	263.824	cl29371	Transposase_21 superfamily
g74	superfamily	9.02E-05	43.6421	cl31762	PTZ00368 superfamily
	superfamily	1.36E-10	58.8852	cl29674	Retrotrans_gag superfamily

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
	superfamily	1.32E-13	67.9029	cl14782	RNase_H_like superfamily
	specific	1.76E-91	288.34	cd01647	RT_LTR
g75					
g76	superfamily	9.12E-03	34.7835	cl29994	Herpes_ICP4_C superfamily
g77					
g78	superfamily	3.90E-03	40.3069	cl33720	PHA03247 superfamily
g79	superfamily	6.34E-28	103.042	cl10012	DnaQ_like_exo superfamily
g80	specific	1.16E-15	67.2967	cd00018	AP2
g81					
g82					
g83	superfamily	1.21E-03	37.8032	cl03830	Transposase_24 superfamily
g84					
g85					
g86	superfamily	1.17E-07	48.8266	cl02808	RT_like superfamily
g87					
g88	superfamily	1.00E-21	100.079	cl31552	PLN03059 superfamily
	superfamily	9.93E-07	48.0592	cl26047	Retrotran_gag_2 superfamily
g89	superfamily	6.59E-04	39.1997	cl26047	Retrotran_gag_2 superfamily
	superfamily	6.32E-14	70.4759	cl06662	RVT_2 superfamily
g90					
g91	superfamily	5.24E-22	91.0149	cl14782	RNase_H_like superfamily
g92	superfamily	4.70E-03	35.9958	cl21462	bZIP superfamily
	superfamily	3.66E-03	40.7118	cl33724	PHA03321 superfamily
g93	superfamily	5.83E-20	84.8517	cl14782	RNase_H_like superfamily

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
	superfamily	5.36E-03	34.9379	cl18014	TSC22 superfamily
g94	superfamily	1.52E-04	43.9454	cl25511	Golgin_A5 superfamily
	superfamily	2.24E-05	47.6257	cl33720	PHA03247 superfamily
g95	superfamily	1.18E-04	45.2924	cl37540	Herpes_BLLF1 superfamily
	superfamily	6.33E-14	70.3169	cl02808	RT_like superfamily
g96	superfamily	3.23E-110	330.997	cl21453	PKc_like superfamily
	superfamily	3.51E-22	101.078	cl33413	PLN00113 superfamily
g97	specific	3.77E-09	53.8796	cd00303	retropepsin_like
g98					
g99	specific	6.18E-04	37.5221	pfam00249	Myb_DNA-binding
	superfamily	1.56E-23	98.5174	cl02808	RT_like superfamily
g100	superfamily	2.52E-38	135.342	cl33561	PLN02758 superfamily
g101					
g102	specific	4.76E-03	32.9999	pfam06747	CHCH
g103	superfamily	1.33E-109	329.456	cl21453	PKc_like superfamily
	superfamily	2.34E-21	98.7668	cl33413	PLN00113 superfamily
g104					
g105					
g106					
g107	specific	9.07E-50	170.806	pfam06584	DIRP
	specific	2.63E-04	39.4481	pfam00249	Myb_DNA-binding
g108					
g109					
g110	superfamily	2.11E-38	144.722	cl12078	p450 superfamily

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g111	specific	9.22E-18	74.638	pfam14223	Retrotran_gag_2
g112					
g113	specific	2.51E-32	121.811	cd09279	RNase_HI_like
	specific	6.41E-50	173.936	cd01647	RT_LTR
	specific	4.56E-12	63.4374	pfam00665	rve
g114	superfamily	9.94E-05	43.4602	cl34364	PspA superfamily
g115					
g116	superfamily	2.60E-07	51.5833	cl00490	EEP superfamily
g117					
g118	specific	4.62E-23	90.7895	pfam13966	zf-RVT
g119	specific	1.67E-04	36.8519	pfam06747	CHCH
g120	specific	1.90E-04	36.4667	pfam06747	CHCH
g121					
g122	superfamily	3.33E-53	184.509	cl12078	p450 superfamily
g123	superfamily	9.94E-11	63.1571	cl06662	RVT_2 superfamily
g124					
g125					
g126	superfamily	4.66E-03	34.4146	cl00828	CbiD superfamily
g127	superfamily	3.51E-22	89.5864	cl29707	PLN02993 superfamily
g128					
g129					
g130					
g131					
g132	specific	4.16E-05	41.0567	pfam00646	F-box

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g133	superfamily	3.19E-13	63.3722	cl17037	NBD_sugar-kinase_HSP70_actin superfamily
	superfamily	1.41E-09	52.9296	cl17169	RRM_SF superfamily
g134					
g135	superfamily	1.55E-32	124.306	cl12078	p450 superfamily
g136					
g137	superfamily	4.07E-03	35.4555	cl07846	DBD_Tnp_Mut superfamily
	superfamily	8.18E-14	66.8388	cl24015	DDE_Tnp_ISL3 superfamily
g137	specific	6.03E-03	36.4398	pfam03247	Prothymosin
g138	superfamily	1.69E-57	198.546	cl36299	PRK13796 superfamily
g139	specific	7.46E-32	111.888	pfam00179	UQ_con
g140	superfamily	1.25E-12	59.4534	cl05007	EBP superfamily
g141	specific	6.48E-21	87.7771	cd00303	retropepsin_like
	superfamily	1.42E-03	38.4696	cl29674	Retrotrans_gag superfamily
	specific	7.33E-34	125.683	cd09274	RNase_HI_RT_Ty3
g142	specific	8.21E-14	72.2355	COG0265	DegQ
g143					
g144					
g145					
g146	superfamily	1.46E-04	44.1957	cl37731	BASP1 superfamily
	superfamily	3.69E-06	50.4905	cl30946	SCP-1 superfamily
	superfamily	3.97E-04	41.9701	cl04444	Transposase_28 superfamily
g147	superfamily	5.26E-4	37.316	cl11403	pepsin_retropepsin_like superfamily
g148	superfamily	2.19E-3	40.6074	cl29689	PHA03309 superfamily
	specific	6.34E-12	61.5816	pfam03732	Retrotrans_gag

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g149	superfamily	4.28E-14	41.1955	cl37069	SMC_prok_B superfamily
g150					
g151	specific	2.37E-12	63.1224	pfam03732	Retrotrans_gag
g152					
g153	superfamily	7.46E-3	34.6176	cl29674	Retrotrans_gag superfamily
g154	specific	7.32E-05	42.3236	cd00303	retropepsin_like
	superfamily	4.26E-10	58.2729	cl14782	RNase_H_like superfamily
	specific	4.29E-34	126.818	cd09279	RNase_HI_like
	specific	2.43E-73	240.19	cd01647	RT_LTR
	specific	1.78E-14	70.371	pfam00665	rve
g155	superfamily	1.23E-06	48.9036	cl04444	Transposase_28 superfamily
g156	specific	3.20E-11	59.2637	cd10017	B3_DNA
g157	superfamily	5.02E-05	44.1589	cl33720	PHA03247 superfamily
g158					
g159					
g160					
g161					
g162	superfamily	5.64E-3	36.1944	cl21584	Tryp_SPc superfamily
g163	specific	1.42E-57	196.631	pfam14214	Helitron_like_N
	specific	1.88E-110	351.285	pfam05970	PIF1
	specific	7.51E-08	50.6379	cd18809	SF1_C_RecD
g164	specific	1.79E-4	40.8168	pfam13365	Trypsin_2
g165					
g166	superfamily	9.19E-14	63.7515	cl10012	DnaQ_like_exo superfamily

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g167	superfamily	2.56E-06	49.1027	cl33662	PLN03210 superfamily
g168	superfamily	2.38E-10	58.1148	cl29674	Retrotrans_gag superfamily
	superfamily	3.22E-18	81.75	cl14782	RNase_H_like superfamily
	superfamily	9.82E-10	58.7609	cl02808	RT_like superfamily
	specific	1.58E-14	70.7562	pfam00665	rve
g169					
g170	superfamily	8.02E-86	259.512	cl21453	PKc_like superfamily
g171					
g172	specific	5.13E-09	53.5405	pfam13976	gag_pre-integrs
	superfamily	5.35E-4	41.1257	cl26047	Retrotran_gag_2 superfamily
	specific	4.93E-72	236.211	cd09272	RNase_HI_RT_Ty1
	specific	4.85E-16	75.3786	pfam00665	rve
	superfamily	8.61E-75	248.053	cl06662	RVT_2 superfamily
g173	superfamily	7.81E-05	43.686	cl34894	AIR1 superfamily
	superfamily	5.86E-3	39.5365	cl33720	PHA03247 superfamily
g174	superfamily	5.97E-3	39.6377	cl35953	rne superfamily
	superfamily	2.84E-08	53.526	cl04444	Transposase_28 superfamily
g175	superfamily	2.28E-3	38.0844	cl29674	Retrotrans_gag superfamily
	superfamily	1.35E-08	53.6505	cl14782	RNase_H_like superfamily
	specific	1.87E-36	133.752	cd09279	RNase_HI_like
	specific	2.43E-73	239.805	cd01647	RT_LTR
g176					
g177					
g178	superfamily	3.76E-3	36.5773	cl04444	Transposase_28 superfamily

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g179	specific	3.61E-07	48.1016	cd00303	retropepsin_like
g180	specific	2.32E-67	230.334	pfam13968	DUF4220
	specific	1.73E-44	164.465	pfam13968	DUF4220
	specific	7.11E-21	86.8209	pfam04578	DUF594
g181	specific	7.11E-12	62.6806	pfam13947	GUB_WAK_bind
	superfamily	1.81E-08	53.0506	cl16494	GUB_WAK_bind superfamily
	specific	1.16E-84	275.306	cd14066	STKc_IRAK
g182	specific	5.84E-73	247.283	pfam13968	DUF4220
	superfamily	5.56E-38	145.975	cl16508	DUF4220 superfamily
	specific	8.40E-22	89.9025	pfam04578	DUF594
	superfamily	2.11E-38	145.108	cl21453	PKc_like superfamily
g183					
g184	superfamily	2.71E-15	72.3314	cl23759	GT1 superfamily
	superfamily	4.20E-10	57.7654	cl27657	Motile_Sperm superfamily
	superfamily	2.36E-4	43.2229	cl23802	Peptidase_C48 superfamily
g185	superfamily	2.14E-3	41.0135	cl33662	PLN03210 superfamily
g186	superfamily	2.64E-3	40.5397	cl33723	PHA03307 superfamily
	superfamily	3.49E-05	46.4075	cl36108	PRK12323 superfamily
g187	specific	5.48E-13	64.699	pfam00635	Motile_Sperm
g188	superfamily	6.45E-08	51.1246	cl16494	GUB_WAK_bind superfamily
	specific	6.69E-92	289.558	cd14066	STKc_IRAK
g189	superfamily	1.28E-20	84.265	cl02808	RT_like superfamily
g190					
g191	superfamily	3.03E-37	142.894	cl16508	DUF4220 superfamily

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
	specific	2.09E-21	87.9765	pfam04578	DUF594
	specific	1.51E-14	68.5633	pfam13976	gag_pre-integrs
	superfamily	9.61E-4	41.586	cl06662	RVT_2 superfamily
g192	specific	9.97E-19	76.5371	pfam13966	zf-RVT
g193	specific	4.09E-81	269.624	pfam13968	DUF4220
	specific	6.47E-47	167.468	cd01650	RT_nLTR_like
g194					
g195	superfamily	8.88E-3	33.9375	cl25898	Atg14 superfamily
g196	superfamily	1.05E-05	44.1329	cl19105	Sina superfamily
g197					
g198					
g199					
g200	specific	1.44E-3	39.9514	COG4886	LRR
	superfamily	1.10E-59	199.014	cl03887	Mlo superfamily
g201					
g202	superfamily	6.04E-4	42.6181	cl33720	PHA03247 superfamily
	superfamily	1.50E-07	49.9852	cl26047	Retrotran_gag_2 superfamily
g203	superfamily	8.65E-07	44.9891	cl16495	DUF4216 superfamily
g204					
g205	specific	1.27E-19	85.0384	pfam14223	Retrotran_gag_2
	superfamily	3.30E-34	130.567	cl06662	RVT_2 superfamily
	specific	9.33E-3	33.956	smart00343	ZnF_C2HC
g206	superfamily	4.97E-09	53.143	cl27657	Motile_Sperm superfamily
g207					

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g208					
g209	superfamily	1.03E-3	41.8477	cl33720	PHA03247 superfamily
	superfamily	6.33E-06	46.9006	cl02808	RT_like superfamily
g210	superfamily	4.02E-62	200.577	cl21453	PKc_like superfamily
g211					
g212	superfamily	1.74E-3	35.5395	cl22382	DUF4219 superfamily
	superfamily	2.46E-12	63.4672	cl26047	Retrotran_gag_2 superfamily
g213	specific	2.17E-05	43.8644	cd00303	retropepsin_like
	superfamily	9.18E-19	83.2908	cl14782	RNase_H_like superfamily
	specific	3.75E-55	189.344	cd01647	RT_LTR
	superfamily	7.16E-07	48.7998	cl21549	rve superfamily
g214	superfamily	5.94E-4	43.1327	cl37070	SMC_prok_A superfamily
	superfamily	2.68E-06	47.7481	cl04444	Transposase_28 superfamily
g215	superfamily	1.06E-06	50.8367	cl37070	SMC_prok_A superfamily
g216	superfamily	3.13E-15	72.5253	cl14782	RNase_H_like superfamily
	specific	2.57E-26	104.477	cd09279	RNase_HI_like
	superfamily	2.01E-11	63.3833	cl02808	RT_like superfamily
g217	superfamily	1.04E-30	108.239	cl29356	PTZ00070 superfamily
g218					
g219					
g220	superfamily	7.01E-3	38.3809	cl33720	PHA03247 superfamily
g221					
g222					
g223					

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g224					
g225					
g226					
g227					
g228	specific	7.25E-3	34.0939	pfam03478	DUF295
g229	superfamily	6.53E-4	40.7643	cl04955	LanC_like superfamily
g230					
g231	superfamily	8.08E-09	55.452	cl04444	Transposase_28 superfamily
g232	specific	2.03E-75	224.397	cd01647	RT_LTR
g233	superfamily	1.80E-17	79.7207	cl06662	RVT_2 superfamily
g234					
g235	superfamily	5.36E-4	39.1423	cl36576	RX-CC_like superfamily
g236					
g237	superfamily	1.89E-32	122.881	cl26397	NB-ARC superfamily
	superfamily	5.11E-4	39.5275	cl36576	RX-CC_like superfamily
g238	specific	3.61E-3	37.3621	pfam13976	gag_pre-integrs
	superfamily	2.16E-3	42.6181	cl33720	PHA03247 superfamily
	superfamily	1.13E-08	57.3791	cl06662	RVT_2 superfamily
	superfamily	1.88E-06	50.4455	cl06662	RVT_2 superfamily
g239					
g240					
g241	specific	2.49E-19	79.2604	pfam14223	Retrotran_gag_2
g242					
g243	superfamily	1.06E-20	86.3925	cl14782	RNase_H_like superfamily

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g244					
g245					
g246					
g247					
g248					
g249					
g250					
g251					
g252	superfamily	7.47E-4	39.7795	cl37892	Retinal superfamily
g253					
g254	superfamily	8.65E-4	43.2406	cl37647	Myosin_tail_1 superfamily
	superfamily	3.17E-3	41.1955	cl37069	SMC_prok_B superfamily
	superfamily	9.00E-06	46.9777	cl04444	Transposase_28 superfamily
g255	superfamily	9.32E-3	36.9308	cl11403	pepsin_retropepsin_like superfamily
	superfamily	2.86E-4	44.9457	cl30053	PHA03291 superfamily
	superfamily	6.26E-05	43.092	cl29674	Retrotrans_gag superfamily
	superfamily	5.37E-15	72.5253	cl14782	RNase_H_like superfamily
	specific	7.47E-44	155.323	cd09279	RNase_HI_like
	specific	1.23E-22	96.5105	cd01647	RT_LTR
g256					
g257	superfamily	5.94E-38	136.33	cl16501	DUF4218 superfamily
	superfamily	6.73E-12	61.1746	cl16504	Transpos_assoc superfamily
	superfamily	1.51E-17	81.625	cl29371	Transposase_21 superfamily
g258					

Continued on next page

Table B.1 – continued from previous page

Predicted Protein	Hit Type	E-value	Bitscore	Accession	Short Name
g259	specific	5.66E-22	88.7359	pfam05920	Homeobox_KN
	superfamily	1.17E-47	163.696	cl02677	POX superfamily
g260	superfamily	6.17E-21	85.8622	cl29371	Transposase_21 superfamily
g261	superfamily	3.33E-3	35.5395	cl22382	DUF4219 superfamily
	specific	5.02E-18	80.8012	pfam14223	Retrotran_gag_2
	specific	2.19E-41	148	cd09272	RNase_HI_RT_Ty1
	superfamily	2.34E-05	46.5935	cl06662	RVT_2 superfamily
	specific	4.47E-06	43.6536	pfam13696	zf-CCHC_2
g262					

Protein domains identified in the AUGUSTUS predicted proteins of the Cadenza *Snn3-B1* pseudomolecule. Protein domains are predicted using NCBI's conserved domain database (Marchler-Bauer et al., 2010).

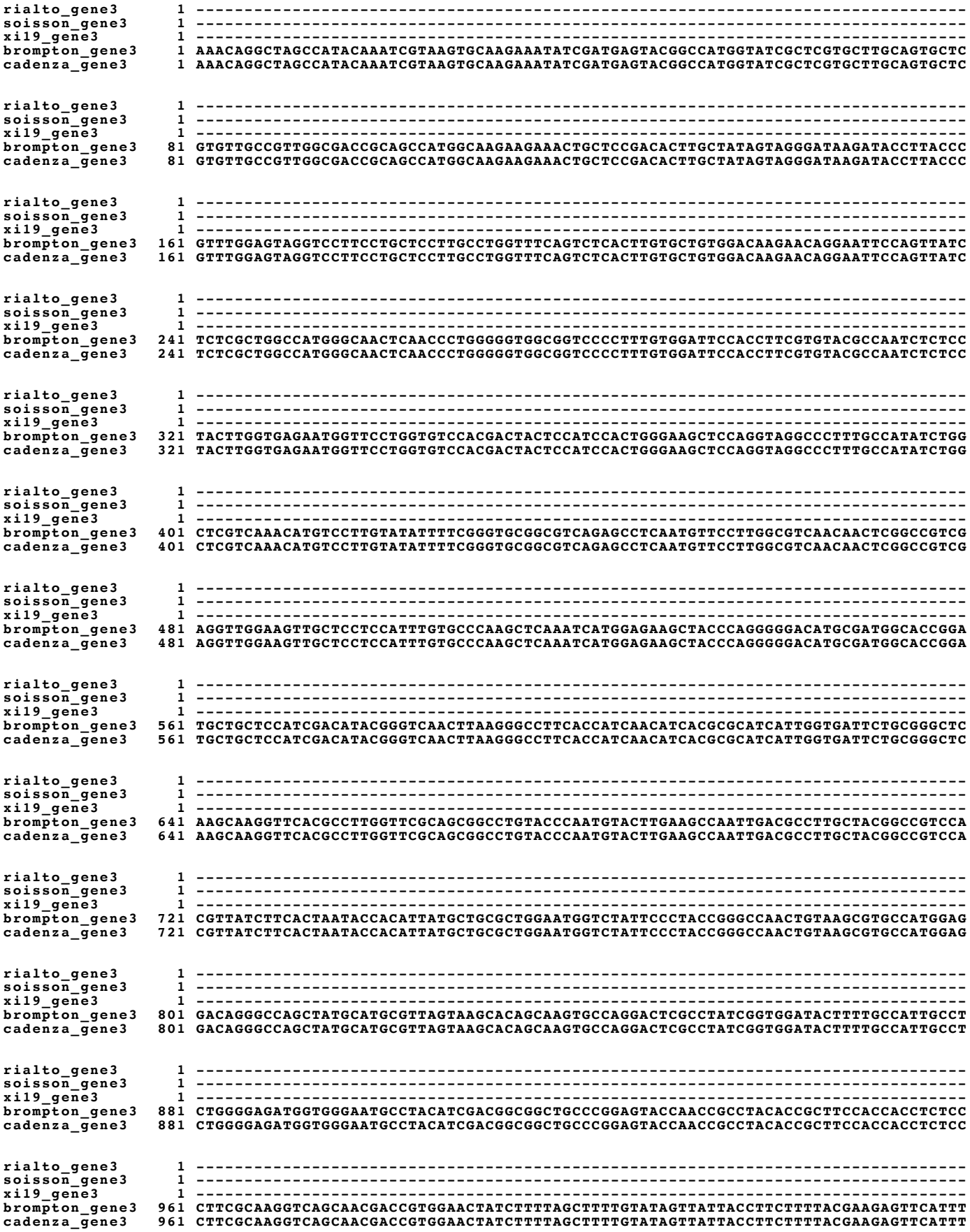
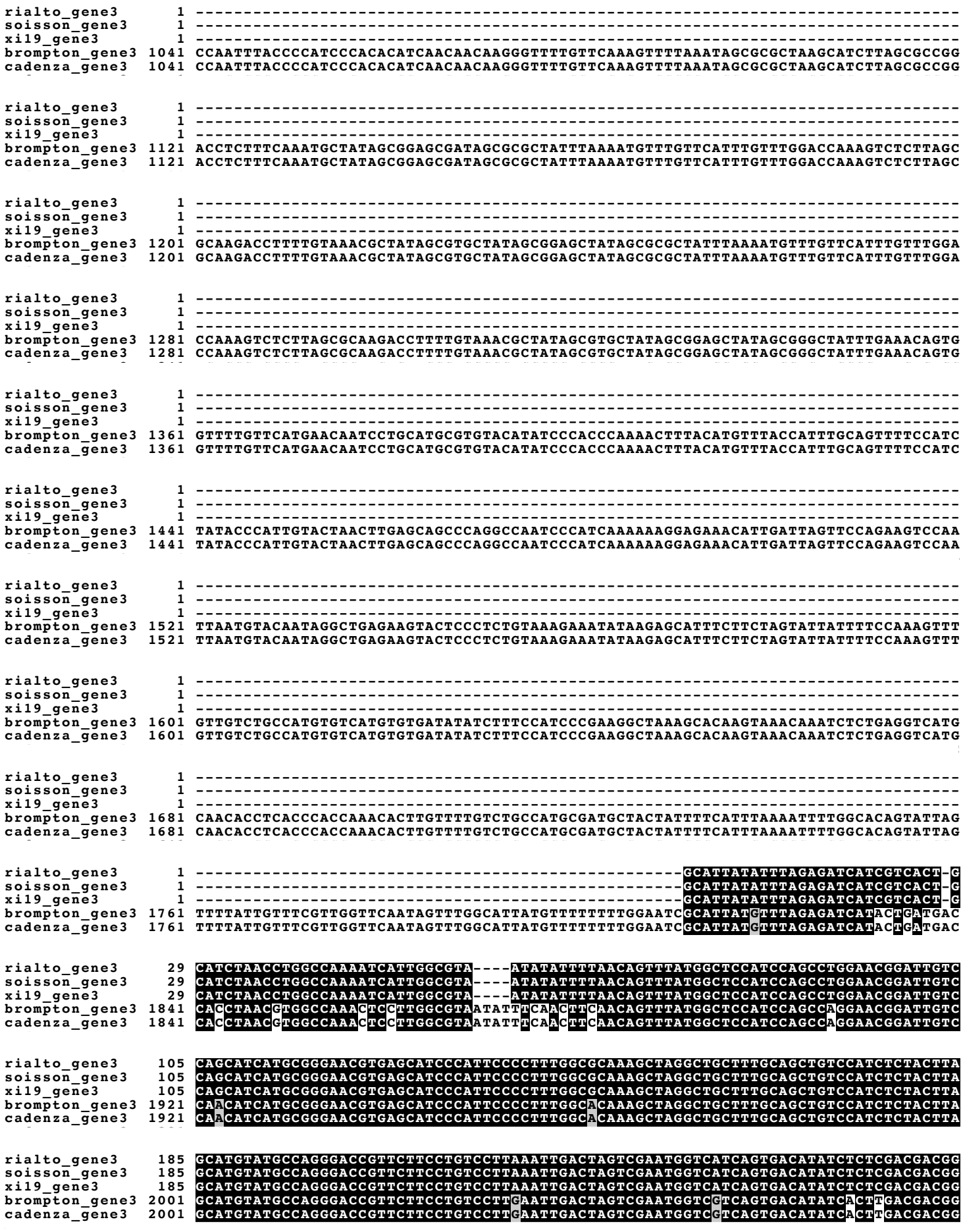
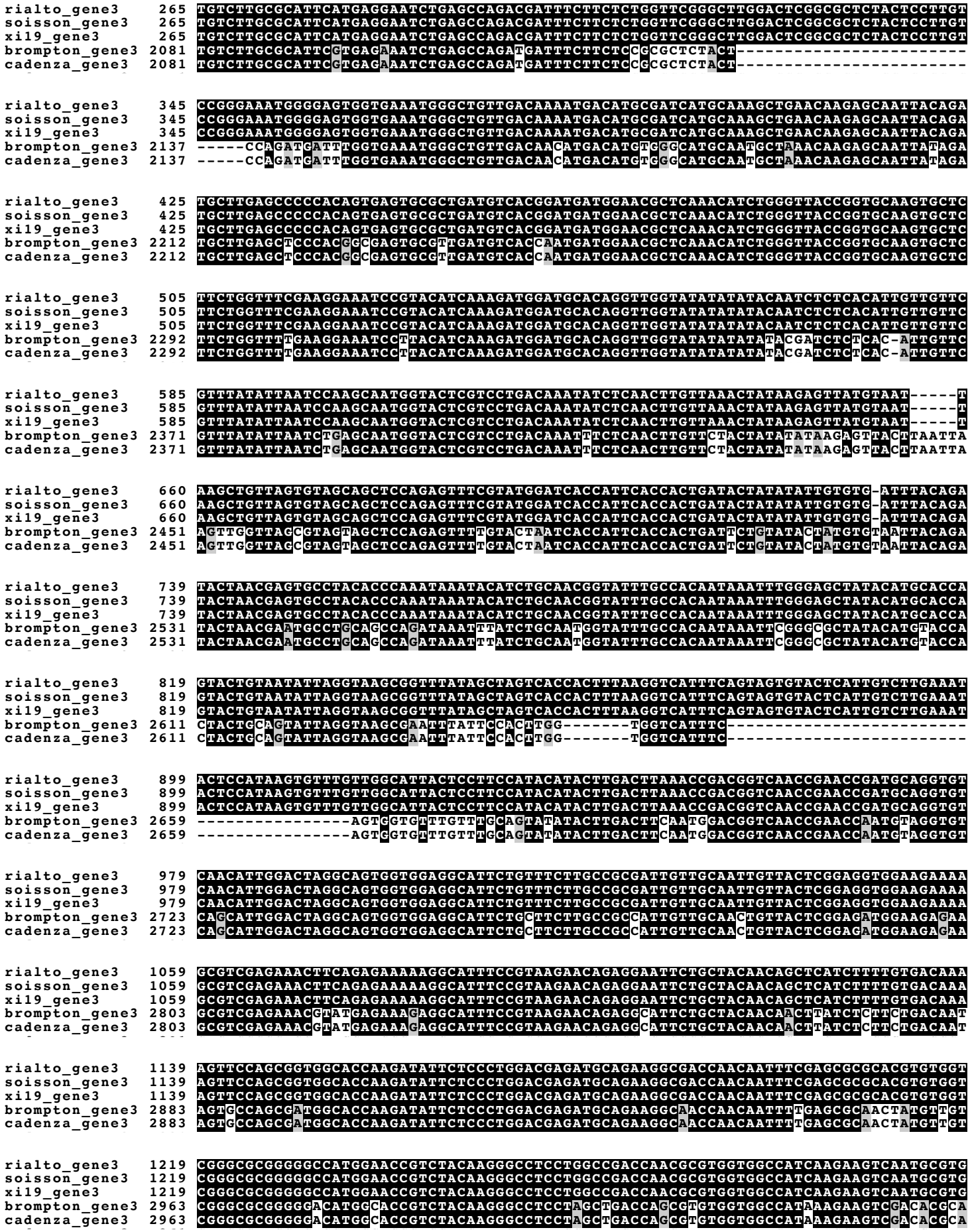
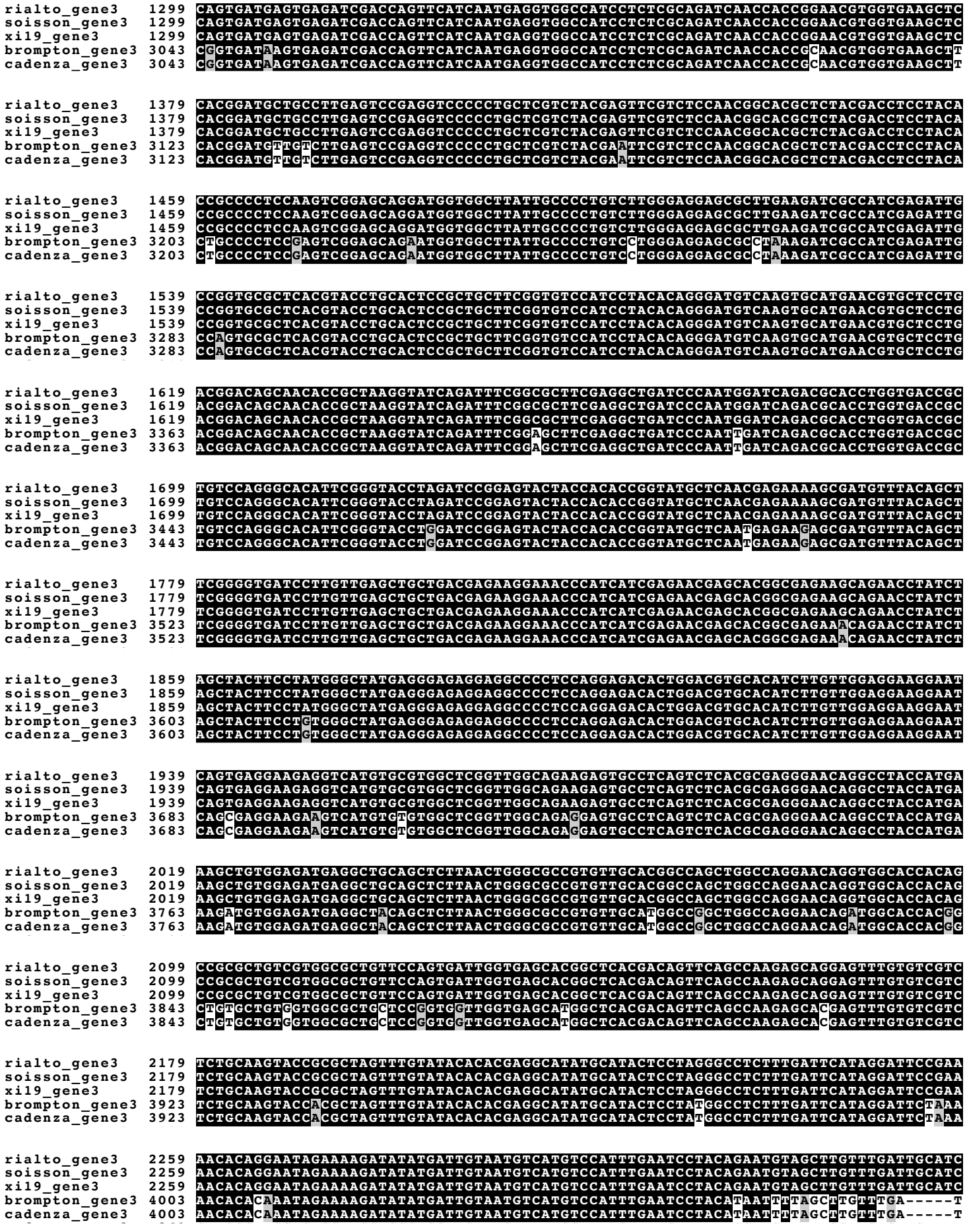
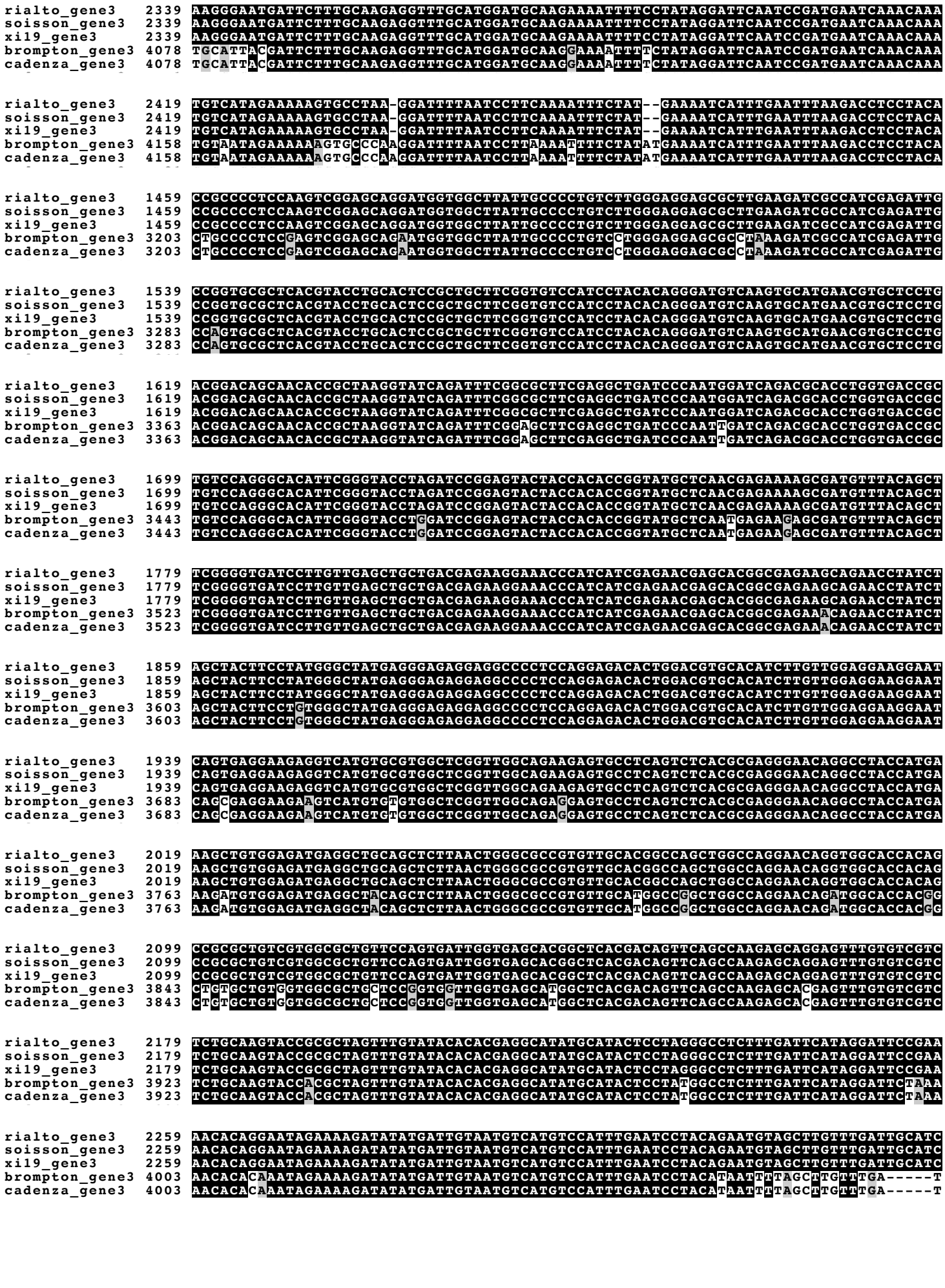


Figure B.1 – Alignment of g188 from the cultivars Rialto, Soissons, Xi19, Brompton and Cadenza.


rialto_gene3 1 ----------
soisson_gene3 1 ----------
xi19_gene3 1 ----------
brompton_gene3 1041 CCAATTTACCCCATCCCACACATCAACAACAAGGGTTTTGTTCAAAGTTTTAAATAGCGCGCTAAGCATCTTAGCGCCGG
cadenza_gene3 1041 CCAATTTACCCCATCCCACACATCAACAACAAGGGTTTTGTTCAAAGTTTTAAATAGCGCGCTAAGCATCTTAGCGCCGG
rialto_gene3 1 ----------
soisson_gene3 1 ----------
xi19_gene3 1 ----------
brompton_gene3 1121 ACCTCTTTCAAATGCTATAGCGGAGCGATAGCGCGCTATTTAAAATGTTTGTTCATTTGTTTGGACCAAAGTCTCTTAGC
cadenza_gene3 1121 ACCTCTTTCAAATGCTATAGCGGAGCGATAGCGCGCTATTTAAAATGTTTGTTCATTTGTTTGGACCAAAGTCTCTTAGC
rialto_gene3 1 ----------
soisson_gene3 1 ----------
xi19_gene3 1 ----------
brompton_gene3 1201 GCAAGACCTTTTGTAAACGCTATAGCGTGCTATAGCGGAGCTATAGCGCGCTATTTAAAATGTTTGTTCATTTGTTTGGA
cadenza_gene3 1201 GCAAGACCTTTTGTAAACGCTATAGCGTGCTATAGCGGAGCTATAGCGCGCTATTTAAAATGTTTGTTCATTTGTTTGGA
rialto_gene3 1 ----------
soisson_gene3 1 ----------
xi19_gene3 1 ----------
brompton_gene3 1281 CCAAAGTCTCTTAGCGCAAGACCTTTTGTAAACGCTATAGCGTGCTATAGCGGAGCTATAGCGGGCTATTTGAAACAGTG
cadenza_gene3 1281 CCAAAGTCTCTTAGCGCAAGACCTTTTGTAAACGCTATAGCGTGCTATAGCGGAGCTATAGCGGGCTATTTGAAACAGTG
rialto_gene3 1 ----------
soisson_gene3 1 ----------
xi19_gene3 1 ----------
brompton_gene3 1361 GTTTTGTTCATGAACAATCCTGCATGCGTGTACATATCCCACCCAAAACTTTACATGTTTACCATTTGCAGTTTTCCATC
cadenza_gene3 1361 GTTTTGTTCATGAACAATCCTGCATGCGTGTACATATCCCACCCAAAACTTTACATGTTTACCATTTGCAGTTTTCCATC
rialto_gene3 1 ----------
soisson_gene3 1 ----------
xi19_gene3 1 ----------
brompton_gene3 1441 TATACCCATTGTACTAACTTGAGCAGCCCAGGCCAATCCCATCAAAAAAGGAGAAACATTGATTAGTTCCAGAAGTCCAA
cadenza_gene3 1441 TATACCCATTGTACTAACTTGAGCAGCCCAGGCCAATCCCATCAAAAAAGGAGAAACATTGATTAGTTCCAGAAGTCCAA
rialto_gene3 1 ----------
soisson_gene3 1 ----------
xi19_gene3 1 ----------
brompton_gene3 1521 TTAATGTACAATAGGCTGAGAAGTACTCCCTCTGTAAAGAAATATAAGAGCATTTCTTCTAGTATTATTTTCCAAAGTTT
cadenza_gene3 1521 TTAATGTACAATAGGCTGAGAAGTACTCCCTCTGTAAAGAAATATAAGAGCATTTCTTCTAGTATTATTTTCCAAAGTTT
rialto_gene3 1 ----------
soisson_gene3 1 ----------
xi19_gene3 1 ----------
brompton_gene3 1601 GTTGTCTGCCATGTGTCATGTGTGATATATCTTTCCATCCCGAAGGCTAAAGCACAAGTAAACAAATCTCTGAGGTCATG
cadenza_gene3 1601 GTTGTCTGCCATGTGTCATGTGTGATATATCTTTCCATCCCGAAGGCTAAAGCACAAGTAAACAAATCTCTGAGGTCATG
rialto_gene3 1 ----------
soisson_gene3 1 ----------
xi19_gene3 1 ----------
brompton_gene3 1681 CAACACCTCACCCACCAAACACTTGTTTTGTCTGCCATGCGATGCTACTATTTTCATTTAAAATTTTGGCACAGTATTAG
cadenza_gene3 1681 CAACACCTCACCCACCAAACACTTGTTTTGTCTGCCATGCGATGCTACTATTTTCATTTAAAATTTTGGCACAGTATTAG
rialto_gene3 1 ----------GCATTATATTTAGAGATCATCGTCACT-G
soisson_gene3 1 ----------GCATTATATTTAGAGATCATCGTCACT-G
xi19_gene3 1 ----------GCATTATATTTAGAGATCATCGTCACT-G
brompton_gene3 1761 TTTTATTGTTTCGTTGGTTCAATAGTTTGGCATTATGTTTTTTTTGGAATCGCATTATGTTTAGAGATCATACTGATGAC
cadenza_gene3 1761 TTTTATTGTTTCGTTGGTTCAATAGTTTGGCATTATGTTTTTTTTGGAATCGCATTATGTTTAGAGATCATACTGATGAC
rialto_gene3 29 CATCTAACCTGGCCAAAATCATTGGCGTA----ATATATTTTAACAGTTTATGGCTCCATCCAGCCTGGAACGGATTGTC
soisson_gene3 29 CATCTAACCTGGCCAAAATCATTGGCGTA----ATATATTTTAACAGTTTATGGCTCCATCCAGCCTGGAACGGATTGTC
xi19_gene3 29 CATCTAACCTGGCCAAAATCATTGGCGTA----ATATATTTTAACAGTTTATGGCTCCATCCAGCCTGGAACGGATTGTC
brompton_gene3 1841 CACCTAACGTGGCCAAACTCCTTGGCGTAATATTTCAACTTCAACAGTTTATGGCTCCATCCAGCCAGGAACGGATTGTC
cadenza_gene3 1841 CACCTAACGTGGCCAAACTCCTTGGCGTAATATTTCAACTTCAACAGTTTATGGCTCCATCCAGCCAGGAACGGATTGTC
rialto_gene3 105 CAGCATCATGCGGGAACGTGAGCATCCCATTCCCCTTTGGCGCAAAGCTAGGCTGCTTTGCAGCTGTCCATCTCTACTTA
soisson_gene3 105 CAGCATCATGCGGGAACGTGAGCATCCCATTCCCCTTTGGCGCAAAGCTAGGCTGCTTTGCAGCTGTCCATCTCTACTTA
xi19_gene3 105 CAGCATCATGCGGGAACGTGAGCATCCCATTCCCCTTTGGCGCAAAGCTAGGCTGCTTTGCAGCTGTCCATCTCTACTTA
brompton_gene3 1921 CAACATCATGCGGGAACGTGAGCATCCCATTCCCCTTTGGCACAAAGCTAGGCTGCTTTGCAGCTGTCCATCTCTACTTA
cadenza_gene3 1921 CAACATCATGCGGGAACGTGAGCATCCCATTCCCCTTTGGCACAAAGCTAGGCTGCTTTGCAGCTGTCCATCTCTACTTA
rialto_gene3 185 GCATGTATGCCAGGGACCGTTCTTCCTGTCCTTAAATTGACTAGTCGAATGGTCATCAGTGACATATCTCTCGACGACGG
soisson_gene3 185 GCATGTATGCCAGGGACCGTTCTTCCTGTCCTTAAATTGACTAGTCGAATGGTCATCAGTGACATATCTCTCGACGACGG
xi19_gene3 185 GCATGTATGCCAGGGACCGTTCTTCCTGTCCTTAAATTGACTAGTCGAATGGTCATCAGTGACATATCTCTCGACGACGG
brompton_gene3 2001 GCATGTATGCCAGGGACCGTTCTTCCTGTCCTTGAATTGACTAGTCGAATGGTCGTCAGTGACATATCACTTGACGACGG
cadenza_gene3 2001 GCATGTATGCCAGGGACCGTTCTTCCTGTCCTTGAATTGACTAGTCGAATGGTCGTCAGTGACATATCACTTGACGACGG


rialto_gene3 265 TGTCTTGCGCATTCATGAGGAATCTGAGCCAGACGATTTCTTCTCTGGTTCGGGCTTGGACTCGGCGCTCTACTCCTTGT
soisson_gene3 265 TGTCTTGCGCATTCATGAGGAATCTGAGCCAGACGATTTCTTCTCTGGTTCGGGCTTGGACTCGGCGCTCTACTCCTTGT
xi19_gene3 265 TGTCTTGCGCATTCATGAGGAATCTGAGCCAGACGATTTCTTCTCTGGTTCGGGCTTGGACTCGGCGCTCTACTCCTTGT
brompton_gene3 2081 TGTCTTGCGCATTCGTGAGAAATCTGAGCCAGATGATTTCTTCTCCGCGCTCTACT----------
cadenza_gene3 2081 TGTCTTGCGCATTCGTGAGAAATCTGAGCCAGATGATTTCTTCTCCGCGCTCTACT----------
rialto_gene3 345 CCGGGAAATGGGGAGTGGTGAAATGGGCTGTTGACAAAATGACATGCGATCATGCAAAGCTGAACAAGAGCAATTACAGA
soisson_gene3 345 CCGGGAAATGGGGAGTGGTGAAATGGGCTGTTGACAAAATGACATGCGATCATGCAAAGCTGAACAAGAGCAATTACAGA
xi19_gene3 345 CCGGGAAATGGGGAGTGGTGAAATGGGCTGTTGACAAAATGACATGCGATCATGCAAAGCTGAACAAGAGCAATTACAGA
brompton_gene3 2137 -----CCAGATGATTTGGTGAAATGGGCTGTTGACAACATGACATGTGGGCATGCAATGCTAAACAAGAGCAATTATAGA
cadenza_gene3 2137 -----CCAGATGATTTGGTGAAATGGGCTGTTGACAACATGACATGTGGGCATGCAATGCTAAACAAGAGCAATTATAGA
rialto_gene3 425 TGCTTGAGCCCCCACAGTGAGTGCGCTGATGTCACGGATGATGGAACGCTCAAACATCTGGGTTACCGGTGCAAGTGCTC
soisson_gene3 425 TGCTTGAGCCCCCACAGTGAGTGCGCTGATGTCACGGATGATGGAACGCTCAAACATCTGGGTTACCGGTGCAAGTGCTC
xi19_gene3 425 TGCTTGAGCCCCCACAGTGAGTGCGCTGATGTCACGGATGATGGAACGCTCAAACATCTGGGTTACCGGTGCAAGTGCTC
brompton_gene3 2212 TGCTTGAGCTCCCACGGCGAGTGCGTTGATGTCACCAATGATGGAACGCTCAAACATCTGGGTTACCGGTGCAAGTGCTC
cadenza_gene3 2212 TGCTTGAGCTCCCACGGCGAGTGCGTTGATGTCACCAATGATGGAACGCTCAAACATCTGGGTTACCGGTGCAAGTGCTC
rialto_gene3 505 TTCTGGTTTCGAAGGAAATCCGTACATCAAAGATGGATGCACAGGTTGGTATATATATACAATCTCTCACATTGTTGTTC
soisson_gene3 505 TTCTGGTTTCGAAGGAAATCCGTACATCAAAGATGGATGCACAGGTTGGTATATATATACAATCTCTCACATTGTTGTTC
xi19_gene3 505 TTCTGGTTTCGAAGGAAATCCGTACATCAAAGATGGATGCACAGGTTGGTATATATATACAATCTCTCACATTGTTGTTC
brompton_gene3 2292 TTCTGGTTTTGAAGGAAATCCTTACATCAAAGATGGATGCACAGGTTGGTATATATATATACGATCTCTCAC-ATTGTTC
cadenza_gene3 2292 TTCTGGTTTTGAAGGAAATCCTTACATCAAAGATGGATGCACAGGTTGGTATATATATATACGATCTCTCAC-ATTGTTC
rialto_gene3 585 GTTTATATTAATCCAAGCAATGGTACTCGTCCTGACAAATATCTCAACTTGTTAAACTATAAGAGTTATGTAAT-----T
soisson_gene3 585 GTTTATATTAATCCAAGCAATGGTACTCGTCCTGACAAATATCTCAACTTGTTAAACTATAAGAGTTATGTAAT-----T
xi19_gene3 585 GTTTATATTAATCCAAGCAATGGTACTCGTCCTGACAAATATCTCAACTTGTTAAACTATAAGAGTTATGTAAT-----T
brompton_gene3 2371 GTTTATATTAATCTGAGCAATGGTACTCGTCCTGACAAATTTCTCAACTTGTTCTACTATATATAAGAGTTACTTAATTA
cadenza_gene3 2371 GTTTATATTAATCTGAGCAATGGTACTCGTCCTGACAAATTTCTCAACTTGTTCTACTATATATAAGAGTTACTTAATTA
rialto_gene3 660 AAGCTGTTAGTGTAGCAGCTCCAGAGTTTCGTATGGATCACCATTCACCACTGATACTATATATTGTGTG-ATTTACAGA
soisson_gene3 660 AAGCTGTTAGTGTAGCAGCTCCAGAGTTTCGTATGGATCACCATTCACCACTGATACTATATATTGTGTG-ATTTACAGA
xi19_gene3 660 AAGCTGTTAGTGTAGCAGCTCCAGAGTTTCGTATGGATCACCATTCACCACTGATACTATATATTGTGTG-ATTTACAGA
brompton_gene3 2451 AGTTGGTTAGCGTAGTAGCTCCAGAGTTTTGTACTAATCACCATTCACCACTGATTCTGTATACTATGTGTAATTACAGA
cadenza_gene3 2451 AGTTGGTTAGCGTAGTAGCTCCAGAGTTTTGTACTAATCACCATTCACCACTGATTCTGTATACTATGTGTAATTACAGA
rialto_gene3 739 TACTAACGAGTGCCTACACCCAAATAAATACATCTGCAACGGTATTTGCCACAATAAATTTGGGAGCTATACATGCACCA
soisson_gene3 739 TACTAACGAGTGCCTACACCCAAATAAATACATCTGCAACGGTATTTGCCACAATAAATTTGGGAGCTATACATGCACCA
xi19_gene3 739 TACTAACGAGTGCCTACACCCAAATAAATACATCTGCAACGGTATTTGCCACAATAAATTTGGGAGCTATACATGCACCA
brompton_gene3 2531 TACTAACGAATGCCTGCAGCCAGATAAATTTATCTGCAATGGTATTTGCCACAATAAATTCGGGCGCTATACATGTACCA
cadenza_gene3 2531 TACTAACGAATGCCTGCAGCCAGATAAATTTATCTGCAATGGTATTTGCCACAATAAATTCGGGCGCTATACATGTACCA
rialto_gene3 819 GTACTGTAATATTAGGTAAGCGGTTTATAGCTAGTCACCACTTTAAGGTCATTTCAGTAGTGTACTCATTGTCTTGAAAT
soisson_gene3 819 GTACTGTAATATTAGGTAAGCGGTTTATAGCTAGTCACCACTTTAAGGTCATTTCAGTAGTGTACTCATTGTCTTGAAAT
xi19_gene3 819 GTACTGTAATATTAGGTAAGCGGTTTATAGCTAGTCACCACTTTAAGGTCATTTCAGTAGTGTACTCATTGTCTTGAAAT
brompton_gene3 2611 CTACTGCAGTATTAGGTAAGCGAATTTATTCCACTTGG-------TGGTCATTTC----------
cadenza_gene3 2611 CTACTGCAGTATTAGGTAAGCGAATTTATTCCACTTGG-------TGGTCATTTC----------
rialto_gene3 899 ACTCCATAAGTGTTTGTTGGCATTACTCCTTCCATACATACTTGACTTAAACCGACGGTCAACCGAACCGATGCAGGTGT
soisson_gene3 899 ACTCCATAAGTGTTTGTTGGCATTACTCCTTCCATACATACTTGACTTAAACCGACGGTCAACCGAACCGATGCAGGTGT
xi19_gene3 899 ACTCCATAAGTGTTTGTTGGCATTACTCCTTCCATACATACTTGACTTAAACCGACGGTCAACCGAACCGATGCAGGTGT
brompton_gene3 2659 ----------------AGTGGTGTTTGTTTGCAGTATATACTTGACTTCAATGGACGGTCAACCGAACCAATGTAGGTGT
cadenza_gene3 2659 ----------------AGTGGTGTTTGTTTGCAGTATATACTTGACTTCAATGGACGGTCAACCGAACCAATGTAGGTGT
rialto_gene3 979 CAACATTGGACTAGGCAGTGGTGGAGGCATTCTGTTTCTTGCCGCGATTGTTGCAATTGTTACTCGGAGGTGGAAGAAAA
soisson_gene3 979 CAACATTGGACTAGGCAGTGGTGGAGGCATTCTGTTTCTTGCCGCGATTGTTGCAATTGTTACTCGGAGGTGGAAGAAAA
xi19_gene3 979 CAACATTGGACTAGGCAGTGGTGGAGGCATTCTGTTTCTTGCCGCGATTGTTGCAATTGTTACTCGGAGGTGGAAGAAAA
brompton_gene3 2723 CAGCATTGGACTAGGCAGTGGTGGAGGCATTCTGCTTCTTGCCGCCATTGTTGCAACTGTTACTCGGAGATGGAAGAGAA
cadenza_gene3 2723 CAGCATTGGACTAGGCAGTGGTGGAGGCATTCTGCTTCTTGCCGCCATTGTTGCAACTGTTACTCGGAGATGGAAGAGAA
rialto_gene3 1059 GCGTCGAGAAACTTCAGAGAAAAAGGCATTTCCGTAAGAACAGAGGAATTCTGCTACAACAGCTCATCTTTTGTGACAAA
soisson_gene3 1059 GCGTCGAGAAACTTCAGAGAAAAAGGCATTTCCGTAAGAACAGAGGAATTCTGCTACAACAGCTCATCTTTTGTGACAAA
xi19_gene3 1059 GCGTCGAGAAACTTCAGAGAAAAAGGCATTTCCGTAAGAACAGAGGAATTCTGCTACAACAGCTCATCTTTTGTGACAAA
brompton_gene3 2803 GCGTCGAGAAACGTATGAGAAAGAGGCATTTCCGTAAGAACAGAGGCATTCTGCTACAACAACTTATCTCTTCTGACAAT
cadenza_gene3 2803 GCGTCGAGAAACGTATGAGAAAGAGGCATTTCCGTAAGAACAGAGGCATTCTGCTACAACAACTTATCTCTTCTGACAAT
rialto_gene3 1139 AGTTCCAGCGGTGGCACCAAGATATTCTCCCTGGACGAGATGCAGAAGGCGACCAACAATTTCGAGCGCGCACGTGTGGT
soisson_gene3 1139 AGTTCCAGCGGTGGCACCAAGATATTCTCCCTGGACGAGATGCAGAAGGCGACCAACAATTTCGAGCGCGCACGTGTGGT
xi19_gene3 1139 AGTTCCAGCGGTGGCACCAAGATATTCTCCCTGGACGAGATGCAGAAGGCGACCAACAATTTCGAGCGCGCACGTGTGGT
brompton_gene3 2883 AGTGCCAGCGATGGCACCAAGATATTCTCCCTGGACGAGATGCAGAAGGCAACCAACAATTTTGAGCGCAACTATGTTGT
cadenza_gene3 2883 AGTGCCAGCGATGGCACCAAGATATTCTCCCTGGACGAGATGCAGAAGGCAACCAACAATTTTGAGCGCAACTATGTTGT
rialto_gene3 1219 CGGGCGCGGGGGGCCATGGAACCGTCTACAAGGGCCTCCTGGCCGACCAACGCGTGGTGGCCATCAAGAAGTCAATGCGTG
soisson_gene3 1219 CGGGCGCGGGGGGCCATGGAACCGTCTACAAGGGCCTCCTGGCCGACCAACGCGTGGTGGCCATCAAGAAGTCAATGCGTG
xi19_gene3 1219 CGGGCGCGGGGGGCCATGGAACCGTCTACAAGGGCCTCCTGGCCGACCAACGCGTGGTGGCCATCAAGAAGTCAATGCGTG
brompton_gene3 2963 CGGGCGCGGGGGGACATGGCACCGTCTACAAGGGCCTCCTAGCTGACCAGCGTGTGGTGGCCATAAAGAAGTCGACACGCA
cadenza_gene3 2963 CGGGCGCGGGGGGACATGGCACCGTCTACAAGGGCCTCCTAGCTGACCAGCGTGTGGTGGCCATAAAGAAGTCGACACGCA


rialto_gene3 1299 CAGTGATGAGTGAGATCGACCAGTTCATCAATGAGGTGGCCATCCTCTCGCAGATCAACCACCGGAACGTGGTGAAGCTC
soisson_gene3 1299 CAGTGATGAGTGAGATCGACCAGTTCATCAATGAGGTGGCCATCCTCTCGCAGATCAACCACCGGAACGTGGTGAAGCTC
xi19_gene3 1299 CAGTGATGAGTGAGATCGACCAGTTCATCAATGAGGTGGCCATCCTCTCGCAGATCAACCACCGGAACGTGGTGAAGCTC
brompton_gene3 3043 CGGTGATAAGTGAGATCGACCAGTTCATCAATGAGGTGGCCATCCTCTCGCAGATCAACCACCGCAACGTGGTGAAGCTT
cadenza_gene3 3043 CGGTGATAAGTGAGATCGACCAGTTCATCAATGAGGTGGCCATCCTCTCGCAGATCAACCACCGCAACGTGGTGAAGCTT
rialto_gene3 1379 CACGGATGCTGCCTTGAGTCCGAGGTCCCCCTGCTCGTCTACGAGTTCGTCTCCAACGGCACGCTCTACGACCTCCTACA
soisson_gene3 1379 CACGGATGCTGCCTTGAGTCCGAGGTCCCCCTGCTCGTCTACGAGTTCGTCTCCAACGGCACGCTCTACGACCTCCTACA
xi19_gene3 1379 CACGGATGCTGCCTTGAGTCCGAGGTCCCCCTGCTCGTCTACGAGTTCGTCTCCAACGGCACGCTCTACGACCTCCTACA
brompton_gene3 3123 CACGGATGTTGTCTTGAGTCCGAGGTCCCCCTGCTCGTCTACGAATTCGTCTCCAACGGCACGCTCTACGACCTCCTACA
cadenza_gene3 3123 CACGGATGTTGTCTTGAGTCCGAGGTCCCCCTGCTCGTCTACGAATTCGTCTCCAACGGCACGCTCTACGACCTCCTACA
rialto_gene3 1459 CCGCCCCTCCAAGTCGGAGCAGGATGGTGGCTTATTGCCCCTGTCTTGGGAGGAGCGCTTGAAGATCGCCATCGAGATTG
soisson_gene3 1459 CCGCCCCTCCAAGTCGGAGCAGGATGGTGGCTTATTGCCCCTGTCTTGGGAGGAGCGCTTGAAGATCGCCATCGAGATTG
xi19_gene3 1459 CCGCCCCTCCAAGTCGGAGCAGGATGGTGGCTTATTGCCCCTGTCTTGGGAGGAGCGCTTGAAGATCGCCATCGAGATTG
brompton_gene3 3203 CTGCCCCTCCGAGTCGGAGCAGAATGGTGGCTTATTGCCCCTGTCCTGGGAGGAGCGCCTAAAGATCGCCATCGAGATTG
cadenza_gene3 3203 CTGCCCCTCCGAGTCGGAGCAGAATGGTGGCTTATTGCCCCTGTCCTGGGAGGAGCGCCTAAAGATCGCCATCGAGATTG
rialto_gene3 1539 CCGGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
soisson_gene3 1539 CCGGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
xi19_gene3 1539 CCGGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
brompton_gene3 3283 CCAGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
cadenza_gene3 3283 CCAGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
rialto_gene3 1619 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGCGCTTCGAGGCTGATCCCAATGGATCAGACGCACCTGGTGACCGC
soisson_gene3 1619 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGCGCTTCGAGGCTGATCCCAATGGATCAGACGCACCTGGTGACCGC
xi19_gene3 1619 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGCGCTTCGAGGCTGATCCCAATGGATCAGACGCACCTGGTGACCGC
brompton_gene3 3363 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGAGCTTCGAGGCTGATCCCAATTGATCAGACGCACCTGGTGACCGC
cadenza_gene3 3363 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGAGCTTCGAGGCTGATCCCAATTGATCAGACGCACCTGGTGACCGC
rialto_gene3 1699 TGTCCAGGGCACATTCGGGTACCTAGATCCGGAGTACTACCACACCGGTATGCTCAACGAGAAAAGCGATGTTTACAGCT
soisson_gene3 1699 TGTCCAGGGCACATTCGGGTACCTAGATCCGGAGTACTACCACACCGGTATGCTCAACGAGAAAAGCGATGTTTACAGCT
xi19_gene3 1699 TGTCCAGGGCACATTCGGGTACCTAGATCCGGAGTACTACCACACCGGTATGCTCAACGAGAAAAGCGATGTTTACAGCT
brompton_gene3 3443 TGTCCAGGGCACATTCGGGTACCTGGATCCGGAGTACTACCACACCGGTATGCTCAATGAGAAGAGCGATGTTTACAGCT
cadenza_gene3 3443 TGTCCAGGGCACATTCGGGTACCTGGATCCGGAGTACTACCACACCGGTATGCTCAATGAGAAGAGCGATGTTTACAGCT
rialto_gene3 1779 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAGCAGAACCTATCT
soisson_gene3 1779 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAGCAGAACCTATCT
xi19_gene3 1779 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAGCAGAACCTATCT
brompton_gene3 3523 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAACAGAACCTATCT
cadenza_gene3 3523 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAACAGAACCTATCT
rialto_gene3 1859 AGCTACTTCCTATGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
soisson_gene3 1859 AGCTACTTCCTATGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
xi19_gene3 1859 AGCTACTTCCTATGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
brompton_gene3 3603 AGCTACTTCCTGTGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
cadenza_gene3 3603 AGCTACTTCCTGTGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
rialto_gene3 1939 CAGTGAGGAAGAGGTCATGTGCGTGGCTCGGTTGGCAGAAGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
soisson_gene3 1939 CAGTGAGGAAGAGGTCATGTGCGTGGCTCGGTTGGCAGAAGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
xi19_gene3 1939 CAGTGAGGAAGAGGTCATGTGCGTGGCTCGGTTGGCAGAAGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
brompton_gene3 3683 CAGCGAGGAAGAAGTCATGTGTGTGGCTCGGTTGGCAGAGGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
cadenza_gene3 3683 CAGCGAGGAAGAAGTCATGTGTGTGGCTCGGTTGGCAGAGGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
rialto_gene3 2019 AAGCTGTGGAGATGAGGCTGCAGCTCTTAACTGGGCGCCGTGTTGCACGGCCAGCTGGCCAGGAACAGGTGGCACCACAG
soisson_gene3 2019 AAGCTGTGGAGATGAGGCTGCAGCTCTTAACTGGGCGCCGTGTTGCACGGCCAGCTGGCCAGGAACAGGTGGCACCACAG
xi19_gene3 2019 AAGCTGTGGAGATGAGGCTGCAGCTCTTAACTGGGCGCCGTGTTGCACGGCCAGCTGGCCAGGAACAGGTGGCACCACAG
brompton_gene3 3763 AAGATGTGGAGATGAGGCTACAGCTCTTAACTGGGCGCCGTGTTGCATGGCCGGCTGGCCAGGAACAGATGGCACCACGG
cadenza_gene3 3763 AAGATGTGGAGATGAGGCTACAGCTCTTAACTGGGCGCCGTGTTGCATGGCCGGCTGGCCAGGAACAGATGGCACCACGG
rialto_gene3 2099 CCGCGCTGTCGTGGCGCTGTTCCAGTGATTGGTGAGCACGGCTCACGACAGTTCAGCCAAGAGCAGGAGTTTGTGTCGTC
soisson_gene3 2099 CCGCGCTGTCGTGGCGCTGTTCCAGTGATTGGTGAGCACGGCTCACGACAGTTCAGCCAAGAGCAGGAGTTTGTGTCGTC
xi19_gene3 2099 CCGCGCTGTCGTGGCGCTGTTCCAGTGATTGGTGAGCACGGCTCACGACAGTTCAGCCAAGAGCAGGAGTTTGTGTCGTC
brompton_gene3 3843 CTGTGCTGTGGTGGCGCTGCTCCGGTGGTTGGTGAGCATGGCTCACGACAGTTCAGCCAAGAGCACGAGTTTGTGTCGTC
cadenza_gene3 3843 CTGTGCTGTGGTGGCGCTGCTCCGGTGGTTGGTGAGCATGGCTCACGACAGTTCAGCCAAGAGCACGAGTTTGTGTCGTC
rialto_gene3 2179 TCTGCAAGTACCGCGCTAGTTTGTATACACACGAGGCATATGCATACTCCTAGGGCCTCTTTGATTCATAGGATTCCGAA
soisson_gene3 2179 TCTGCAAGTACCGCGCTAGTTTGTATACACACGAGGCATATGCATACTCCTAGGGCCTCTTTGATTCATAGGATTCCGAA
xi19_gene3 2179 TCTGCAAGTACCGCGCTAGTTTGTATACACACGAGGCATATGCATACTCCTAGGGCCTCTTTGATTCATAGGATTCCGAA
brompton_gene3 3923 TCTGCAAGTACCACGCTAGTTTGTATACACACGAGGCATATGCATACTCCTATGGCCTCTTTGATTCATAGGATTCTAAA
cadenza_gene3 3923 TCTGCAAGTACCACGCTAGTTTGTATACACACGAGGCATATGCATACTCCTATGGCCTCTTTGATTCATAGGATTCTAAA
rialto_gene3 2259 AACACAGGAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACAGAATGTAGCTTGTTTGATTGCATC
soisson_gene3 2259 AACACAGGAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACAGAATGTAGCTTGTTTGATTGCATC
xi19_gene3 2259 AACACAGGAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACAGAATGTAGCTTGTTTGATTGCATC
brompton_gene3 4003 AACACACAAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACATAATTTTAGCTTGTTTGA-----T
cadenza_gene3 4003 AACACACAAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACATAATTTTAGCTTGTTTGA-----T


rialto_gene3 2339 AAGGGAATGATTCTTTGCAAGAGGTTTGCATGGATGCAAGAAAATTTTCCTATAGGATTCAATCCGATGAATCAAACAAA
soisson_gene3 2339 AAGGGAATGATTCTTTGCAAGAGGTTTGCATGGATGCAAGAAAATTTTCCTATAGGATTCAATCCGATGAATCAAACAAA
xi19_gene3 2339 AAGGGAATGATTCTTTGCAAGAGGTTTGCATGGATGCAAGAAAATTTTCCTATAGGATTCAATCCGATGAATCAAACAAA
brompton_gene3 4078 TGCATTACGATTCTTTGCAAGAGGTTTGCATGGATGCAAGGAAAATTTTCTATAGGATTCAATCCGATGAATCAAACAAA
cadenza_gene3 4078 TGCATTACGATTCTTTGCAAGAGGTTTGCATGGATGCAAGGAAAATTTTCTATAGGATTCAATCCGATGAATCAAACAAA
rialto_gene3 2419 TGTCATAGAAAAAGTGCCTAA-GGATTTTAATCCTTCAAAATTTCTAT--GAAAATCATTTGAATTTAAGACCTCCTACA
soisson_gene3 2419 TGTCATAGAAAAAGTGCCTAA-GGATTTTAATCCTTCAAAATTTCTAT--GAAAATCATTTGAATTTAAGACCTCCTACA
xi19_gene3 2419 TGTCATAGAAAAAGTGCCTAA-GGATTTTAATCCTTCAAAATTTCTAT--GAAAATCATTTGAATTTAAGACCTCCTACA
brompton_gene3 4158 TGTAATAGAAAAAAGTGCCCAAGGATTTTAATCCTTAAAATTTTCTATATGAAAATCATTTGAATTTAAGACCTCCTACA
cadenza_gene3 4158 TGTAATAGAAAAAAGTGCCCAAGGATTTTAATCCTTAAAATTTTCTATATGAAAATCATTTGAATTTAAGACCTCCTACA
rialto_gene3 1459 CCGCCCCTCCAAGTCGGAGCAGGATGGTGGCTTATTGCCCCTGTCTTGGGAGGAGCGCTTGAAGATCGCCATCGAGATTG
soisson_gene3 1459 CCGCCCCTCCAAGTCGGAGCAGGATGGTGGCTTATTGCCCCTGTCTTGGGAGGAGCGCTTGAAGATCGCCATCGAGATTG
xi19_gene3 1459 CCGCCCCTCCAAGTCGGAGCAGGATGGTGGCTTATTGCCCCTGTCTTGGGAGGAGCGCTTGAAGATCGCCATCGAGATTG
brompton_gene3 3203 CTGCCCCTCCGAGTCGGAGCAGAATGGTGGCTTATTGCCCCTGTCCTGGGAGGAGCGCCTAAAGATCGCCATCGAGATTG
cadenza_gene3 3203 CTGCCCCTCCGAGTCGGAGCAGAATGGTGGCTTATTGCCCCTGTCCTGGGAGGAGCGCCTAAAGATCGCCATCGAGATTG
rialto_gene3 1539 CCGGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
soisson_gene3 1539 CCGGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
xi19_gene3 1539 CCGGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
brompton_gene3 3283 CCAGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
cadenza_gene3 3283 CCAGTGCGCTCACGTACCTGCACTCCGCTGCTTCGGTGTCCATCCTACACAGGGATGTCAAGTGCATGAACGTGCTCCTG
rialto_gene3 1619 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGCGCTTCGAGGCTGATCCCAATGGATCAGACGCACCTGGTGACCGC
soisson_gene3 1619 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGCGCTTCGAGGCTGATCCCAATGGATCAGACGCACCTGGTGACCGC
xi19_gene3 1619 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGCGCTTCGAGGCTGATCCCAATGGATCAGACGCACCTGGTGACCGC
brompton_gene3 3363 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGAGCTTCGAGGCTGATCCCAATTGATCAGACGCACCTGGTGACCGC
cadenza_gene3 3363 ACGGACAGCAACACCGCTAAGGTATCAGATTTCGGAGCTTCGAGGCTGATCCCAATTGATCAGACGCACCTGGTGACCGC
rialto_gene3 1699 TGTCCAGGGCACATTCGGGTACCTAGATCCGGAGTACTACCACACCGGTATGCTCAACGAGAAAAGCGATGTTTACAGCT
soisson_gene3 1699 TGTCCAGGGCACATTCGGGTACCTAGATCCGGAGTACTACCACACCGGTATGCTCAACGAGAAAAGCGATGTTTACAGCT
xi19_gene3 1699 TGTCCAGGGCACATTCGGGTACCTAGATCCGGAGTACTACCACACCGGTATGCTCAACGAGAAAAGCGATGTTTACAGCT
brompton_gene3 3443 TGTCCAGGGCACATTCGGGTACCTGGATCCGGAGTACTACCACACCGGTATGCTCAATGAGAAGAGCGATGTTTACAGCT
cadenza_gene3 3443 TGTCCAGGGCACATTCGGGTACCTGGATCCGGAGTACTACCACACCGGTATGCTCAATGAGAAGAGCGATGTTTACAGCT
rialto_gene3 1779 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAGCAGAACCTATCT
soisson_gene3 1779 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAGCAGAACCTATCT
xi19_gene3 1779 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAGCAGAACCTATCT
brompton_gene3 3523 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAACAGAACCTATCT
cadenza_gene3 3523 TCGGGGTGATCCTTGTTGAGCTGCTGACGAGAAGGAAACCCATCATCGAGAACGAGCACGGCGAGAAACAGAACCTATCT
rialto_gene3 1859 AGCTACTTCCTATGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
soisson_gene3 1859 AGCTACTTCCTATGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
xi19_gene3 1859 AGCTACTTCCTATGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
brompton_gene3 3603 AGCTACTTCCTGTGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
cadenza_gene3 3603 AGCTACTTCCTGTGGGCTATGAGGGAGAGGAGGCCCCTCCAGGAGACACTGGACGTGCACATCTTGTTGGAGGAAGGAAT
rialto_gene3 1939 CAGTGAGGAAGAGGTCATGTGCGTGGCTCGGTTGGCAGAAGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
soisson_gene3 1939 CAGTGAGGAAGAGGTCATGTGCGTGGCTCGGTTGGCAGAAGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
xi19_gene3 1939 CAGTGAGGAAGAGGTCATGTGCGTGGCTCGGTTGGCAGAAGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
brompton_gene3 3683 CAGCGAGGAAGAAGTCATGTGTGTGGCTCGGTTGGCAGAGGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
cadenza_gene3 3683 CAGCGAGGAAGAAGTCATGTGTGTGGCTCGGTTGGCAGAGGAGTGCCTCAGTCTCACGCGAGGGAACAGGCCTACCATGA
rialto_gene3 2019 AAGCTGTGGAGATGAGGCTGCAGCTCTTAACTGGGCGCCGTGTTGCACGGCCAGCTGGCCAGGAACAGGTGGCACCACAG
soisson_gene3 2019 AAGCTGTGGAGATGAGGCTGCAGCTCTTAACTGGGCGCCGTGTTGCACGGCCAGCTGGCCAGGAACAGGTGGCACCACAG
xi19_gene3 2019 AAGCTGTGGAGATGAGGCTGCAGCTCTTAACTGGGCGCCGTGTTGCACGGCCAGCTGGCCAGGAACAGGTGGCACCACAG
brompton_gene3 3763 AAGATGTGGAGATGAGGCTACAGCTCTTAACTGGGCGCCGTGTTGCATGGCCGGCTGGCCAGGAACAGATGGCACCACGG
cadenza_gene3 3763 AAGATGTGGAGATGAGGCTACAGCTCTTAACTGGGCGCCGTGTTGCATGGCCGGCTGGCCAGGAACAGATGGCACCACGG
rialto_gene3 2099 CCGCGCTGTCGTGGCGCTGTTCCAGTGATTGGTGAGCACGGCTCACGACAGTTCAGCCAAGAGCAGGAGTTTGTGTCGTC
soisson_gene3 2099 CCGCGCTGTCGTGGCGCTGTTCCAGTGATTGGTGAGCACGGCTCACGACAGTTCAGCCAAGAGCAGGAGTTTGTGTCGTC
xi19_gene3 2099 CCGCGCTGTCGTGGCGCTGTTCCAGTGATTGGTGAGCACGGCTCACGACAGTTCAGCCAAGAGCAGGAGTTTGTGTCGTC
brompton_gene3 3843 CTGTGCTGTGGTGGCGCTGCTCCGGTGGTTGGTGAGCATGGCTCACGACAGTTCAGCCAAGAGCACGAGTTTGTGTCGTC
cadenza_gene3 3843 CTGTGCTGTGGTGGCGCTGCTCCGGTGGTTGGTGAGCATGGCTCACGACAGTTCAGCCAAGAGCACGAGTTTGTGTCGTC
rialto_gene3 2179 TCTGCAAGTACCGCGCTAGTTTGTATACACACGAGGCATATGCATACTCCTAGGGCCTCTTTGATTCATAGGATTCCGAA
soisson_gene3 2179 TCTGCAAGTACCGCGCTAGTTTGTATACACACGAGGCATATGCATACTCCTAGGGCCTCTTTGATTCATAGGATTCCGAA
xi19_gene3 2179 TCTGCAAGTACCGCGCTAGTTTGTATACACACGAGGCATATGCATACTCCTAGGGCCTCTTTGATTCATAGGATTCCGAA
brompton_gene3 3923 TCTGCAAGTACCACGCTAGTTTGTATACACACGAGGCATATGCATACTCCTATGGCCTCTTTGATTCATAGGATTCTAAA
cadenza_gene3 3923 TCTGCAAGTACCACGCTAGTTTGTATACACACGAGGCATATGCATACTCCTATGGCCTCTTTGATTCATAGGATTCTAAA
rialto_gene3 2259 AACACAGGAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACAGAATGTAGCTTGTTTGATTGCATC
soisson_gene3 2259 AACACAGGAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACAGAATGTAGCTTGTTTGATTGCATC
xi19_gene3 2259 AACACAGGAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACAGAATGTAGCTTGTTTGATTGCATC
brompton_gene3 4003 AACACACAAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACATAATTTTAGCTTGTTTGA-----T
cadenza_gene3 4003 AACACACAAATAGAAAAGATATATGATTGTAATGTCATGTCCATTTGAATCCTACATAATTTTAGCTTGTTTGA-----T

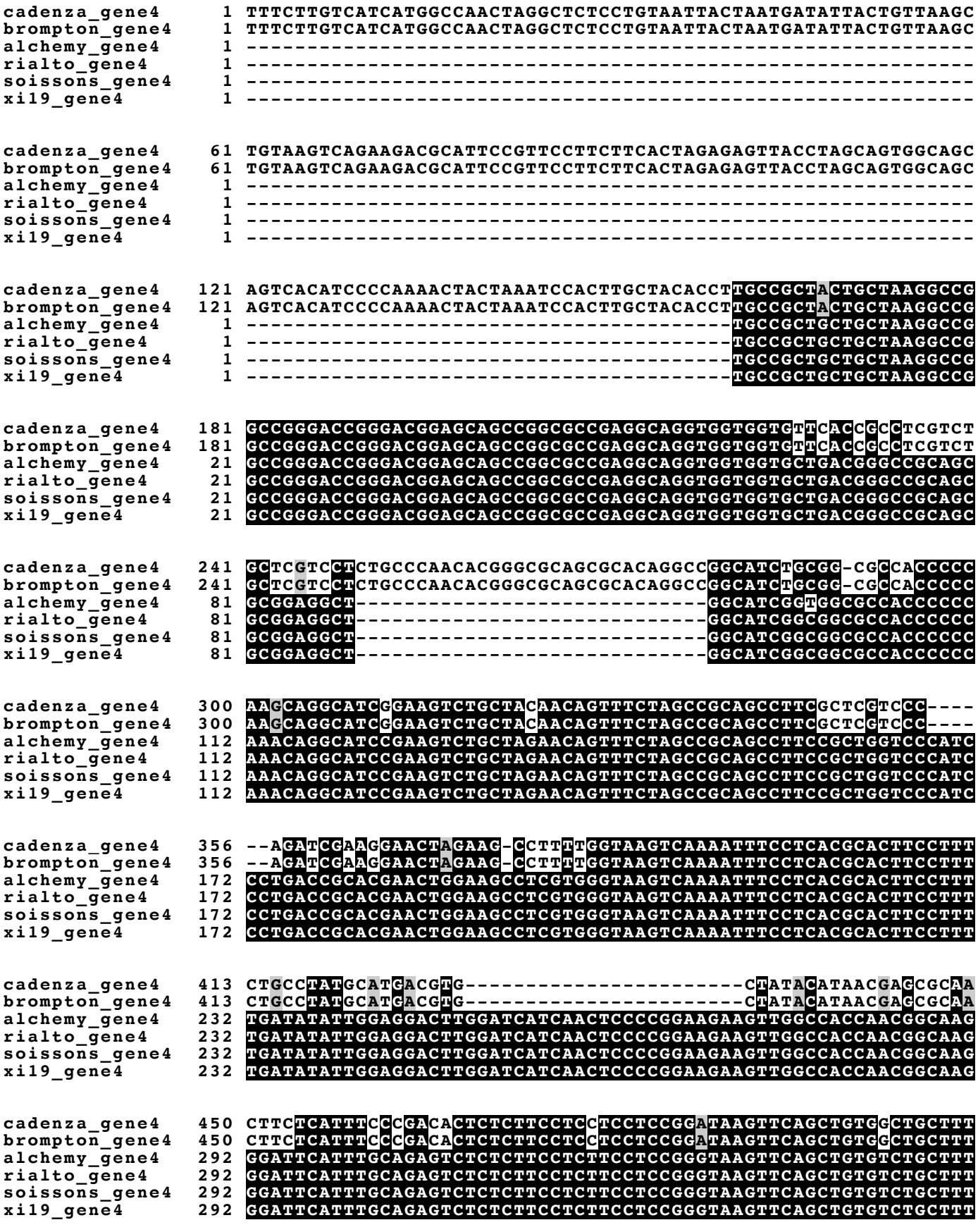
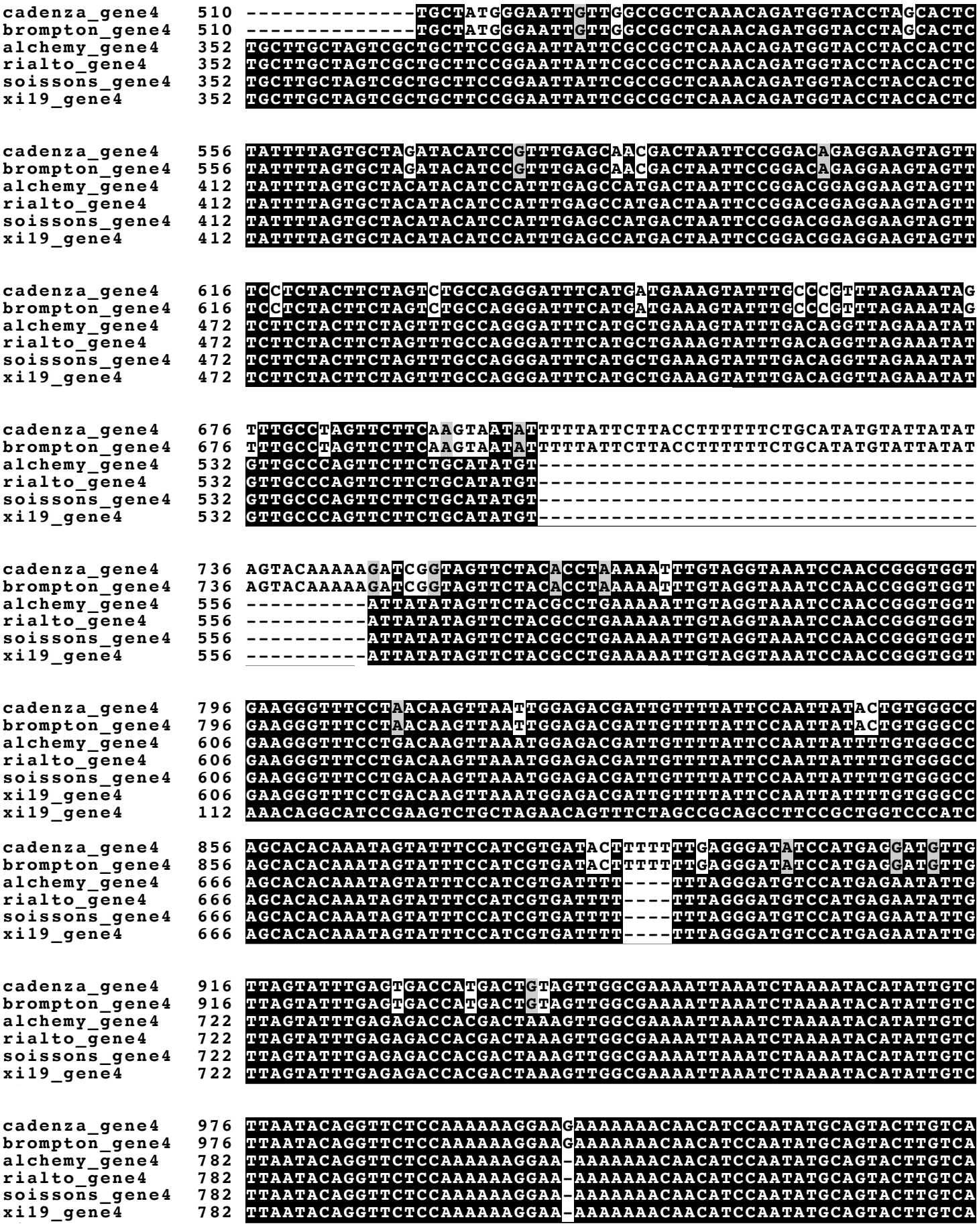
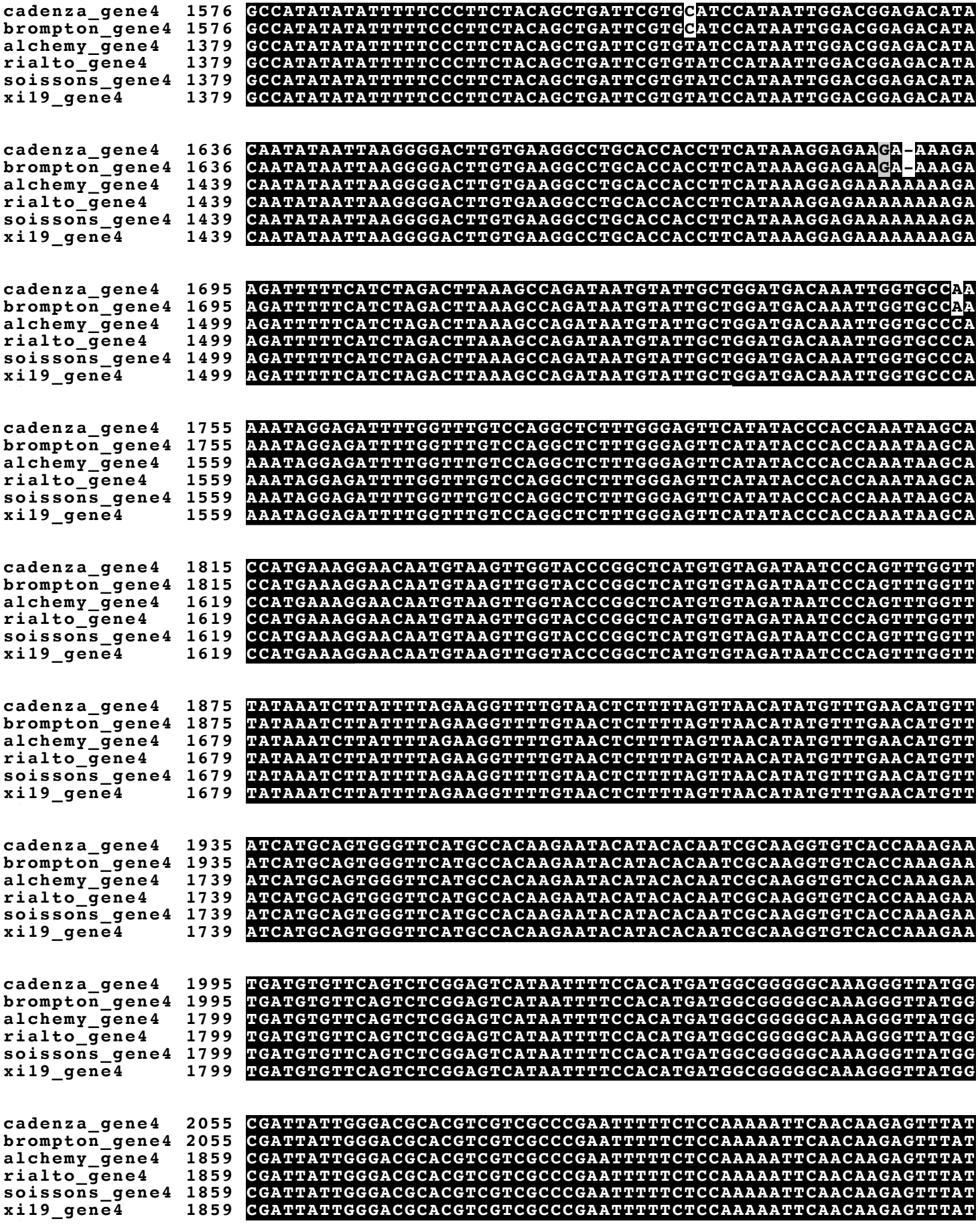


Figure B.2 – Alignment of g187 from the cultivars Cadenza, Brompton, Alchemy, Rialto, Soissons and Xi19.


cadenza_gene4 510 ----------TGCTATGGGAATTGTTGGCCGCTCAAACAGATGGTACCTAGCACTC
brompton_gene4 510 ----------TGCTATGGGAATTGTTGGCCGCTCAAACAGATGGTACCTAGCACTC
alchemy_gene4 352 TGCTTGCTAGTCGCTGCTTCCGGAATTATTCGCCGCTCAAACAGATGGTACCTACCACTC
rialto_gene4 352 TGCTTGCTAGTCGCTGCTTCCGGAATTATTCGCCGCTCAAACAGATGGTACCTACCACTC
soissons_gene4 352 TGCTTGCTAGTCGCTGCTTCCGGAATTATTCGCCGCTCAAACAGATGGTACCTACCACTC
xi19_gene4 352 TGCTTGCTAGTCGCTGCTTCCGGAATTATTCGCCGCTCAAACAGATGGTACCTACCACTC
cadenza_gene4 556 TATTTTAGTGCTAGATACATCCGTTTGAGCAACGACTAATTCCGGACAGAGGAAGTAGTT
brompton_gene4 556 TATTTTAGTGCTAGATACATCCGTTTGAGCAACGACTAATTCCGGACAGAGGAAGTAGTT
alchemy_gene4 412 TATTTTAGTGCTACATACATCCATTTGAGCCATGACTAATTCCGGACGGAGGAAGTAGTT
rialto_gene4 412 TATTTTAGTGCTACATACATCCATTTGAGCCATGACTAATTCCGGACGGAGGAAGTAGTT
soissons_gene4 412 TATTTTAGTGCTACATACATCCATTTGAGCCATGACTAATTCCGGACGGAGGAAGTAGTT
xi19_gene4 412 TATTTTAGTGCTACATACATCCATTTGAGCCATGACTAATTCCGGACGGAGGAAGTAGTT
cadenza_gene4 616 TCCTCTACTTCTAGTCTGCCAGGGATTTCATGATGAAAGTATTTGCCCGTTTAGAAATAG
brompton_gene4 616 TCCTCTACTTCTAGTCTGCCAGGGATTTCATGATGAAAGTATTTGCCCGTTTAGAAATAG
alchemy_gene4 472 TCTTCTACTTCTAGTTTGCCAGGGATTTCATGCTGAAAGTATTTGACAGGTTAGAAATAT
rialto_gene4 472 TCTTCTACTTCTAGTTTGCCAGGGATTTCATGCTGAAAGTATTTGACAGGTTAGAAATAT
soissons_gene4 472 TCTTCTACTTCTAGTTTGCCAGGGATTTCATGCTGAAAGTATTTGACAGGTTAGAAATAT
xi19_gene4 472 TCTTCTACTTCTAGTTTGCCAGGGATTTCATGCTGAAAGTATTTGACAGGTTAGAAATAT
cadenza_gene4 676 TTTGCCTAGTTCTTCAAGTAATATTTTTATTCTTACCTTTTTTCTGCATATGTATTATAT
brompton_gene4 676 TTTGCCTAGTTCTTCAAGTAATATTTTTATTCTTACCTTTTTTCTGCATATGTATTATAT
alchemy_gene4 532 GTTGCCCAGTTCTTCTGCATATGT----------
rialto_gene4 532 GTTGCCCAGTTCTTCTGCATATGT----------
soissons_gene4 532 GTTGCCCAGTTCTTCTGCATATGT----------
xi19_gene4 532 GTTGCCCAGTTCTTCTGCATATGT----------
cadenza_gene4 736 AGTACAAAAAGATCGGTAGTTCTACACCTAAAAATTTGTAGGTAAATCCAACCGGGTGGT
brompton_gene4 736 AGTACAAAAAGATCGGTAGTTCTACACCTAAAAATTTGTAGGTAAATCCAACCGGGTGGT
alchemy_gene4 556 ----------ATTATATAGTTCTACGCCTGAAAAATTGTAGGTAAATCCAACCGGGTGGT
rialto_gene4 556 ----------ATTATATAGTTCTACGCCTGAAAAATTGTAGGTAAATCCAACCGGGTGGT
soissons_gene4 556 ----------ATTATATAGTTCTACGCCTGAAAAATTGTAGGTAAATCCAACCGGGTGGT
xi19_gene4 556 ----------ATTATATAGTTCTACGCCTGAAAAATTGTAGGTAAATCCAACCGGGTGGT
cadenza_gene4 796 GAAGGGTTTCCTAACAAGTTAATTGGAGACGATTGTTTTATTCCAATTATACTGTGGGCC
brompton_gene4 796 GAAGGGTTTCCTAACAAGTTAATTGGAGACGATTGTTTTATTCCAATTATACTGTGGGCC
alchemy_gene4 606 GAAGGGTTTCCTGACAAGTTAAATGGAGACGATTGTTTTATTCCAATTATTTTGTGGGCC
rialto_gene4 606 GAAGGGTTTCCTGACAAGTTAAATGGAGACGATTGTTTTATTCCAATTATTTTGTGGGCC
soissons_gene4 606 GAAGGGTTTCCTGACAAGTTAAATGGAGACGATTGTTTTATTCCAATTATTTTGTGGGCC
xi19_gene4 606 GAAGGGTTTCCTGACAAGTTAAATGGAGACGATTGTTTTATTCCAATTATTTTGTGGGCC
xi19_gene4 112 AAACAGGCATCCGAAGTCTGCTAGAACAGTTTCTAGCCGCAGCCTTCCGCTGGTCCCATC
cadenza_gene4 856 AGCACACAAATAGTATTTCCATCGTGATACTTTTTTTGAGGGATATCCATGAGGATGTTG
brompton_gene4 856 AGCACACAAATAGTATTTCCATCGTGATACTTTTTTTGAGGGATATCCATGAGGATGTTG
alchemy_gene4 666 AGCACACAAATAGTATTTCCATCGTGATTTT----TTTAGGGATGTCCATGAGAATATTG
rialto_gene4 666 AGCACACAAATAGTATTTCCATCGTGATTTT----TTTAGGGATGTCCATGAGAATATTG
soissons_gene4 666 AGCACACAAATAGTATTTCCATCGTGATTTT----TTTAGGGATGTCCATGAGAATATTG
xi19_gene4 666 AGCACACAAATAGTATTTCCATCGTGATTTT----TTTAGGGATGTCCATGAGAATATTG
cadenza_gene4 916 TTAGTATTTGAGTGACCATGACTGTAGTTGGCGAAAATTAAATCTAAAATACATATTGTC
brompton_gene4 916 TTAGTATTTGAGTGACCATGACTGTAGTTGGCGAAAATTAAATCTAAAATACATATTGTC
alchemy_gene4 722 TTAGTATTTGAGAGACCACGACTAAAGTTGGCGAAAATTAAATCTAAAATACATATTGTC
rialto_gene4 722 TTAGTATTTGAGAGACCACGACTAAAGTTGGCGAAAATTAAATCTAAAATACATATTGTC
soissons_gene4 722 TTAGTATTTGAGAGACCACGACTAAAGTTGGCGAAAATTAAATCTAAAATACATATTGTC
xi19_gene4 722 TTAGTATTTGAGAGACCACGACTAAAGTTGGCGAAAATTAAATCTAAAATACATATTGTC
cadenza_gene4 976 TTAATACAGGTTCTCCAAAAAAGGAAGAAAAAAACAACATCCAATATGCAGTACTTGTCA
brompton_gene4 976 TTAATACAGGTTCTCCAAAAAAGGAAGAAAAAAACAACATCCAATATGCAGTACTTGTCA
alchemy_gene4 782 TTAATACAGGTTCTCCAAAAAAGGAA-AAAAAAACAACATCCAATATGCAGTACTTGTCA
rialto_gene4 782 TTAATACAGGTTCTCCAAAAAAGGAA-AAAAAAACAACATCCAATATGCAGTACTTGTCA
soissons_gene4 782 TTAATACAGGTTCTCCAAAAAAGGAA-AAAAAAACAACATCCAATATGCAGTACTTGTCA
xi19_gene4 782 TTAATACAGGTTCTCCAAAAAAGGAA-AAAAAAACAACATCCAATATGCAGTACTTGTCA

```
cadenza_gene4  1036 GATAAAATGTCCTCGTCTGATCCACAATATATCAAATTATGTTTGTTACAAGCAATCACA
brompton_gene4 1036 GATAAAATGTCCTCGTCTGATCCACAATATATCAAATTATGTTTGTTACAAGCAATCACA
alchemy_gene4   841 GATAAAATGTCCTCGTCCGATCCACAATATATCAAATTATGTTTGTTACAAGCAATCACA
rialto_gene4    841 GATAAAATGTCCTCGTCCGATCCACAATATATCAAATTATGTTTGTTACAAGCAATCACA
soissons_gene4  841 GATAAAATGTCCTCGTCCGATCCACAATATATCAAATTATGTTTGTTACAAGCAATCACA
xi19_gene4      841 GATAAAATGTCCTCGTCCGATCCACAATATATCAAATTATGTTTGTTACAAGCAATCACA

cadenza_gene4  1096 GAGGAGTTTTCAGAGAAGATGAAAATTGGTACTGGTGGCTATGGGGAAGTTTACAAGGTA
brompton_gene4 1096 GAGGAGTTTTCAGAGAAGATGAAAATTGGTACTGGTGGCTATGGGGAAGTTTACAAGGTA
alchemy_gene4   901 GAGGAGTTTTCGGAGAAGATGAAAATTGGTACTGGTGGCTATGGGGAAGTTTACAAGGTA
rialto_gene4    901 GAGGAGTTTTCGGAGAAGATGAAAATTGGTACTGGTGGCTATGGGGAAGTTTACAAGGTA
soissons_gene4  901 GAGGAGTTTTCGGAGAAGATGAAAATTGGTACTGGTGGCTATGGGGAAGTTTACAAGGTA
xi19_gene4      901 GAGGAGTTTTCGGAGAAGATGAAAATTGGTACTGGTGGCTATGGGGAAGTTTACAAGGTA

cadenza_gene4  1156 GGTATATGGTCAATCATTTATGTTTGTTATGCTAACTAGTCGTATTATGTTTTCTGTAAT
brompton_gene4 1156 GGTATATGGTCAATCATTTATGTTTGTTATGCTAACTAGTCGTATTATGTTTTCTGTAAT
alchemy_gene4   961 GGT--ATGGTCAATCATTTATGTTTGTTATACTAACTAATCGTATTATGTTTTCTGTAAT
rialto_gene4    961 GGT--ATGGTCAATCATTTATGTTTGTTATACTAACTAATCGTATTATGTTTTCTGTAAT
soissons_gene4  961 GGT--ATGGTCAATCATTTATGTTTGTTATACTAACTAATCGTATTATGTTTTCTGTAAT
xi19_gene4      961 GGT--ATGGTCAATCATTTATGTTTGTTATACTAACTAATCGTATTATGTTTTCTGTAAT

cadenza_gene4  1216 GATCATATGGTTAATATGAACAGGGTGAGCTTAATGGGGACGAAATTGCTGTCAAGAAGC
brompton_gene4 1216 GATCATATGGTTAATATGAACAGGGTGAGCTTAATGGGGACGAAATTGCTGTCAAGAAGC
alchemy_gene4  1019 TATCATATGGTCAATATGAACAGGGAGAGCTTAATGGGGACGAGATTGCTGTCAAGAAGC
rialto_gene4   1019 TATCATATGGTCAATATGAACAGGGAGAGCTTAATGGGGACGAGATTGCTGTCAAGAAGC
soissons_gene4 1019 TATCATATGGTCAATATGAACAGGGAGAGCTTAATGGGGACGAGATTGCTGTCAAGAAGC
xi19_gene4     1019 TATCATATGGTCAATATGAACAGGGAGAGCTTAATGGGGACGAGATTGCTGTCAAGAAGC

cadenza_gene4  1276 TCTTTCCCATCCAAGGAGTTAATGATGAGTCATTTGATAATGAATTCCGTAACCTTAAGA
brompton_gene4 1276 TCTTTCCCATCCAAGGAGTTAATGATGAGTCATTTGATAATGAATTCCGTAACCTTAAGA
alchemy_gene4  1079 TTTTTCCCGTCCAAGGAGTTAATGATGAGTCATTTGATAATGAATTCCGTAACCTTAAGA
rialto_gene4   1079 TTTTTCCCGTCCAAGGAGTTAATGATGAGTCATTTGATAATGAATTCCGTAACCTTAAGA
soissons_gene4 1079 TTTTTCCCGTCCAAGGAGTTAATGATGAGTCATTTGATAATGAATTCCGTAACCTTAAGA
xi19_gene4     1079 TTTTTCCCGTCCAAGGAGTTAATGATGAGTCATTTGATAATGAATTCCGTAACCTTAAGA

cadenza_gene4  1336 AGGTTCGGCATAAAAATGTCATACGGATGATTGGTTACTGCTATGAGACATCACATAGAG
brompton_gene4 1336 AGGTTCGGCATAAAAATGTCATACGGATGATTGGTTACTGCTATGAGACATCACATAGAG
alchemy_gene4  1139 AGGTTCGGCATAAAAATGTCGTACGGATGATTGGTTACTGCTATGAGACATCACATAGAG
rialto_gene4   1139 AGGTTCGGCATAAAAATGTCGTACGGATGATTGGTTACTGCTATGAGACATCACATAGAG
soissons_gene4 1139 AGGTTCGGCATAAAAATGTCGTACGGATGATTGGTTACTGCTATGAGACATCACATAGAG
xi19_gene4     1139 AGGTTCGGCATAAAAATGTCGTACGGATGATTGGTTACTGCTATGAGACATCACATAGAG

cadenza_gene4  1396 ATGTTGAATACAAAGGTCAATTAGTTTGGTCTCAAGTAATAAACAGAGCTCTCTGCTTTG
brompton_gene4 1396 ATGTTGAATACAAAGGTCAATTAGTTTGGTCTCAAGTAATAAACAGAGCTCTCTGCTTTG
alchemy_gene4  1199 ATGTTGAATACAAAGGTCAATTAGTTTGGTCTCAAGTAATAAACAGAGCTCTCTGCTTTG
rialto_gene4   1199 ATGTTGAATACAAAGGTCAATTAGTTTGGTCTCAAGTAATAAACAGAGCTCTCTGCTTTG
soissons_gene4 1199 ATGTTGAATACAAAGGTCAATTAGTTTGGTCTCAAGTAATAAACAGAGCTCTCTGCTTTG
xi19_gene4     1199 ATGTTGAATACAAAGGTCAATTAGTTTGGTCTCAAGTAATAAACAGAGCTCTCTGCTTTG

cadenza_gene4  1456 AGTATATGAAGGGAGGAAGCCTTGCCAAGCATATTTCTGGTATGATGGTACTCTACCTCA
brompton_gene4 1456 AGTATATGAAGGGAGGAAGCCTTGCCAAGCATATTTCTGGTATGATGGTACTCTACCTCA
alchemy_gene4  1259 AGTATATGAAGGGAGGAAGCCTTGCCAAGCATATTTCTGGTATGATGGTACTCTACCTCA
rialto_gene4   1259 AGTATATGAAGGGAGGAAGCCTTGCCAAGCATATTTCTGGTATGATGGTACTCTACCTCA
soissons_gene4 1259 AGTATATGAAGGGAGGAAGCCTTGCCAAGCATATTTCTGGTATGATGGTACTCTACCTCA
xi19_gene4     1259 AGTATATGAAGGGAGGAAGCCTTGCCAAGCATATTTCTGGTATGATGGTACTCTACCTCA

cadenza_gene4  1516 ATTCTGCTATGCTACCATTTACCTTGCAATGAAGATATGATTTTGCCATGTTATGTGTAT
brompton_gene4 1516 ATTCTGCTATGCTACCATTTACCTTGCAATGAAGATATGATTTTGCCATGTTATGTGTAT
alchemy_gene4  1319 ATTCTGCTATGCTACCATTTACCTTGCAATGAAGATATGATTTTGCCATGTTATGTGTAT
rialto_gene4   1319 ATTCTGCTATGCTACCATTTACCTTGCAATGAAGATATGATTTTGCCATGTTATGTGTAT
soissons_gene4 1319 ATTCTGCTATGCTACCATTTACCTTGCAATGAAGATATGATTTTGCCATGTTATGTGTAT
xi19_gene4     1319 ATTCTGCTATGCTACCATTTACCTTGCAATGAAGATATGATTTTGCCATGTTATGTGTAT
```


cadenza_gene4 1576 GCCATATATATTTTTCCCTTCTACAGCTGATTCGTGCATCCATAATTGGACGGAGACATA
brompton_gene4 1576 GCCATATATATTTTTCCCTTCTACAGCTGATTCGTGCATCCATAATTGGACGGAGACATA
alchemy_gene4 1379 GCCATATATATTTTTCCCTTCTACAGCTGATTCGTGTATCCATAATTGGACGGAGACATA
rialto_gene4 1379 GCCATATATATTTTTCCCTTCTACAGCTGATTCGTGTATCCATAATTGGACGGAGACATA
soissons_gene4 1379 GCCATATATATTTTTCCCTTCTACAGCTGATTCGTGTATCCATAATTGGACGGAGACATA
xi19_gene4 1379 GCCATATATATTTTTCCCTTCTACAGCTGATTCGTGTATCCATAATTGGACGGAGACATA
cadenza_gene4 1636 CAATATAATTAAGGGGACTTGTGAAGGCCTGCACCACCTTCATAAAGGAGAAGA-AAAGA
brompton_gene4 1636 CAATATAATTAAGGGGACTTGTGAAGGCCTGCACCACCTTCATAAAGGAGAAGA-AAAGA
alchemy_gene4 1439 CAATATAATTAAGGGGACTTGTGAAGGCCTGCACCACCTTCATAAAGGAGAAAAAAAAGA
rialto_gene4 1439 CAATATAATTAAGGGGACTTGTGAAGGCCTGCACCACCTTCATAAAGGAGAAAAAAAAGA
soissons_gene4 1439 CAATATAATTAAGGGGACTTGTGAAGGCCTGCACCACCTTCATAAAGGAGAAAAAAAAGA
xi19_gene4 1439 CAATATAATTAAGGGGACTTGTGAAGGCCTGCACCACCTTCATAAAGGAGAAAAAAAAGA
cadenza_gene4 1695 AGATTTTTCATCTAGACTTAAAGCCAGATAATGTATTGCTGGATGACAAATTGGTGCCAA
brompton_gene4 1695 AGATTTTTCATCTAGACTTAAAGCCAGATAATGTATTGCTGGATGACAAATTGGTGCCAA
alchemy_gene4 1499 AGATTTTTCATCTAGACTTAAAGCCAGATAATGTATTGCTGGATGACAAATTGGTGCCCA
rialto_gene4 1499 AGATTTTTCATCTAGACTTAAAGCCAGATAATGTATTGCTGGATGACAAATTGGTGCCCA
soissons_gene4 1499 AGATTTTTCATCTAGACTTAAAGCCAGATAATGTATTGCTGGATGACAAATTGGTGCCCA
xi19_gene4 1499 AGATTTTTCATCTAGACTTAAAGCCAGATAATGTATTGCTGGATGACAAATTGGTGCCCA
cadenza_gene4 1755 AAATAGGAGATTTTGGTTTGTCCAGGCTCTTTGGGAGTTCATATACCCACCAAATAAGCA
brompton_gene4 1755 AAATAGGAGATTTTGGTTTGTCCAGGCTCTTTGGGAGTTCATATACCCACCAAATAAGCA
alchemy_gene4 1559 AAATAGGAGATTTTGGTTTGTCCAGGCTCTTTGGGAGTTCATATACCCACCAAATAAGCA
rialto_gene4 1559 AAATAGGAGATTTTGGTTTGTCCAGGCTCTTTGGGAGTTCATATACCCACCAAATAAGCA
soissons_gene4 1559 AAATAGGAGATTTTGGTTTGTCCAGGCTCTTTGGGAGTTCATATACCCACCAAATAAGCA
xi19_gene4 1559 AAATAGGAGATTTTGGTTTGTCCAGGCTCTTTGGGAGTTCATATACCCACCAAATAAGCA
cadenza_gene4 1815 CCATGAAAGGAACAATGTAAGTTGGTACCCGGCTCATGTGTAGATAATCCCAGTTTGGTT
brompton_gene4 1815 CCATGAAAGGAACAATGTAAGTTGGTACCCGGCTCATGTGTAGATAATCCCAGTTTGGTT
alchemy_gene4 1619 CCATGAAAGGAACAATGTAAGTTGGTACCCGGCTCATGTGTAGATAATCCCAGTTTGGTT
rialto_gene4 1619 CCATGAAAGGAACAATGTAAGTTGGTACCCGGCTCATGTGTAGATAATCCCAGTTTGGTT
soissons_gene4 1619 CCATGAAAGGAACAATGTAAGTTGGTACCCGGCTCATGTGTAGATAATCCCAGTTTGGTT
xi19_gene4 1619 CCATGAAAGGAACAATGTAAGTTGGTACCCGGCTCATGTGTAGATAATCCCAGTTTGGTT
cadenza_gene4 1875 TATAAATCTTATTTTAGAAGGTTTTGTAACTCTTTTAGTTAACATATGTTTGAACATGTT
brompton_gene4 1875 TATAAATCTTATTTTAGAAGGTTTTGTAACTCTTTTAGTTAACATATGTTTGAACATGTT
alchemy_gene4 1679 TATAAATCTTATTTTAGAAGGTTTTGTAACTCTTTTAGTTAACATATGTTTGAACATGTT
rialto_gene4 1679 TATAAATCTTATTTTAGAAGGTTTTGTAACTCTTTTAGTTAACATATGTTTGAACATGTT
soissons_gene4 1679 TATAAATCTTATTTTAGAAGGTTTTGTAACTCTTTTAGTTAACATATGTTTGAACATGTT
xi19_gene4 1679 TATAAATCTTATTTTAGAAGGTTTTGTAACTCTTTTAGTTAACATATGTTTGAACATGTT
cadenza_gene4 1935 ATCATGCAGTGGGTTCATGCCACAAGAATACATACACAATCGCAAGGTGTCACCAAAGAA
brompton_gene4 1935 ATCATGCAGTGGGTTCATGCCACAAGAATACATACACAATCGCAAGGTGTCACCAAAGAA
alchemy_gene4 1739 ATCATGCAGTGGGTTCATGCCACAAGAATACATACACAATCGCAAGGTGTCACCAAAGAA
rialto_gene4 1739 ATCATGCAGTGGGTTCATGCCACAAGAATACATACACAATCGCAAGGTGTCACCAAAGAA
soissons_gene4 1739 ATCATGCAGTGGGTTCATGCCACAAGAATACATACACAATCGCAAGGTGTCACCAAAGAA
xi19_gene4 1739 ATCATGCAGTGGGTTCATGCCACAAGAATACATACACAATCGCAAGGTGTCACCAAAGAA
cadenza_gene4 1995 TGATGTGTTCAGTCTCGGAGTCATAATTTTCCACATGATGGCGGGGGCAAAGGGTTATGG
brompton_gene4 1995 TGATGTGTTCAGTCTCGGAGTCATAATTTTCCACATGATGGCGGGGGCAAAGGGTTATGG
alchemy_gene4 1799 TGATGTGTTCAGTCTCGGAGTCATAATTTTCCACATGATGGCGGGGGCAAAGGGTTATGG
rialto_gene4 1799 TGATGTGTTCAGTCTCGGAGTCATAATTTTCCACATGATGGCGGGGGCAAAGGGTTATGG
soissons_gene4 1799 TGATGTGTTCAGTCTCGGAGTCATAATTTTCCACATGATGGCGGGGGCAAAGGGTTATGG
xi19_gene4 1799 TGATGTGTTCAGTCTCGGAGTCATAATTTTCCACATGATGGCGGGGGCAAAGGGTTATGG
cadenza_gene4 2055 CGATTATTGGGACGCACGTCGTCGCCCGAATTTTTCTCCAAAAATTCAACAAGAGTTTAT
brompton_gene4 2055 CGATTATTGGGACGCACGTCGTCGCCCGAATTTTTCTCCAAAAATTCAACAAGAGTTTAT
alchemy_gene4 1859 CGATTATTGGGACGCACGTCGTCGCCCGAATTTTTCTCCAAAAATTCAACAAGAGTTTAT
rialto_gene4 1859 CGATTATTGGGACGCACGTCGTCGCCCGAATTTTTCTCCAAAAATTCAACAAGAGTTTAT
soissons_gene4 1859 CGATTATTGGGACGCACGTCGTCGCCCGAATTTTTCTCCAAAAATTCAACAAGAGTTTAT
xi19_gene4 1859 CGATTATTGGGACGCACGTCGTCGCCCGAATTTTTCTCCAAAAATTCAACAAGAGTTTAT

```
cadenza_gene4  2115 GGAGAGTGTAAGATAGTTTATTTGTCATTTTTACGTTTTAAGGGGTTGTTGTTCCTTATC
brompton_gene4 2115 GGAGAGTGTAAGATAGTTTATTTGTCATTTTTACGTTTTAAGGGGTTGTTGTTCCTTATC
alchemy_gene4  1919 GGAGAGTGTAAGATAGTTTATTTGTCATTTTTACGTTTTAAGGGGTTGTTGTTCCTTATC
rialto_gene4   1919 GGAGAGTGTAAGATAGTTTATTTGTCATTTTTACGTTTTAAGGGGTTGTTGTTCCTTATC
soissons_gene4 1919 GGAGAGTGTAAGATAGTTTATTTGTCATTTTTACGTTTTAAGGGGTTGTTGTTCCTTATC
xi19_gene4     1919 GGAGAGTGTAAGATAGTTTATTTGTCATTTTTACGTTTTAAGGGGTTGTTGTTCCTTATC

cadenza_gene4  2175 TCTGTTTTTCATGCGATTGGTAGATGCACAATATATGTTGCAACTTGACAGGTACAAATA
brompton_gene4 2175 TCTGTTTTTCATGCGATTGGTAGATGCACAATATATGTTGCAACTTGACAGGTACAAATA
alchemy_gene4  1979 TCTGTTTTTCATGCGATTGGTAGATGCACAATATATGTTGCAACTTGACAGGTACAAATA
rialto_gene4   1979 TCTGTTTTTCATGCGATTGGTAGATGCACAATATATGTTGCAACTTGACAGGTACAAATA
soissons_gene4 1979 TCTGTTTTTCATGCGATTGGTAGATGCACAATATATGTTGCAACTTGACAGGTACAAATA
xi19_gene4     1979 TCTGTTTTTCATGCGATTGGTAGATGCACAATATATGTTGCAACTTGACAGGTACAAATA

cadenza_gene4  2235 TACTGGAGGAAAAAGATGCAGGCAACTGAAGATTATAGATGGGACGAAACCGATCTACTA
brompton_gene4 2235 TACTGGAGGAAAAAGATGCAGGCAACTGAAGATTATAGATGGGACGAAACCGATCTACTA
alchemy_gene4  2039 TACTGGAGGAAAAAGATGCAGGCAACTGAAGATTATAGATGGGACGAAACCGATCTACTA
rialto_gene4   2039 TACTGGAGGAAAAAGATGCAGGCAACTGAAGATTATAGATGGGACGAAACCGATCTACTA
soissons_gene4 2039 TACTGGAGGAAAAAGATGCAGGCAACTGAAGATTATAGATGGGACGAAACCGATCTACTA
xi19_gene4     2039 TACTGGAGGAAAAAGATGCAGGCAACTGAAGATTATAGATGGGACGAAACCGATCTACTA

cadenza_gene4  2295 GGAGTAACGAAATGTGTTGATATGGCAATGCGTTGTGTGGAGGACGACAGAGACAATAGA
brompton_gene4 2295 GGAGTAACGAAATGTGTTGATATGGCAATGCGTTGTGTGGAGGACGACAGAGACAATAGA
alchemy_gene4  2099 GGAGTAACGAAATGTGTTGATATGGCAATGCGTTGTGTGGAGGACGACAGAGACAATAGA
rialto_gene4   2099 GGAGTAACGAAATGTGTTGATATGGCAATGCGTTGTGTGGAGGACGACAGAGACAATAGA
soissons_gene4 2099 GGAGTAACGAAATGTGTTGATATGGCAATGCGTTGTGTGGAGGACGACAGAGACAATAGA
xi19_gene4     2099 GGAGTAACGAAATGTGTTGATATGGCAATGCGTTGTGTGGAGGACGACAGAGACAATAGA

cadenza_gene4  2355 CCTTCTACAGAGGACATTATCAGTGAGCTCAAGGAACTAGACTCAAAGATAGACGAGATG
brompton_gene4 2355 CCTTCTACAGAGGACATTATCAGTGAGCTCAAGGAACTAGACTCAAAGATAGACGAGATG
alchemy_gene4  2159 CCTTCTACAGAGGACATTATCAGTGAGCTCAAGGAACTAGACTCAAAGATAGACGAGATG
rialto_gene4   2159 CCTTCTACAGAGGACATTATCAGTGAGCTCAAGGAACTAGACTCAAAGATAGACGAGATG
soissons_gene4 2159 CCTTCTACAGAGGACATTATCAGTGAGCTCAAGGAACTAGACTCAAAGATAGACGAGATG
xi19_gene4     2159 CCTTCTACAGAGGACATTATCAGTGAGCTCAAGGAACTAGACTCAAAGATAGACGAGATG

cadenza_gene4  2415 TTAAAGAAAGATCCTAAACCTCTTACAGGCCAGCTGGTACAATATATATTCATACTCATA
brompton_gene4 2415 TTAAAGAAAGATCCTAAACCTCTTACAGGCCAGCTGGTACAATATATATTCATACTCATA
alchemy_gene4  2219 TTAAAGAAAGATCCTAAACCTCTTACAGGCCAGCTGGTACAATATATATTCATACTCATA
rialto_gene4   2219 TTAAAGAAAGATCCTAAACCTCTTACAGGCCAGCTGGTACAATATATATTCATACTCATA
soissons_gene4 2219 TTAAAGAAAGATCCTAAACCTCTTACAGGCCAGCTGGTACAATATATATTCATACTCATA
xi19_gene4     2219 TTAAAGAAAGATCCTAAACCTCTTACAGGCCAGCTGGTACAATATATATTCATACTCATA

cadenza_gene4  2475 TATGTGGATCCAGCAATGTATAGTTTCTGAACACTAGCTAGTTAACACGTTGAACTTTTG
brompton_gene4 2475 TATGTGGATCCAGCAATGTATAGTTTCTGAACACTAGCTAGTTAACACGTTGAACTTTTG
alchemy_gene4  2279 TATGTGGATCCAGCAATGTATAGTTTCTGAACACTAGCTAGTTAACACGTTGAACTTTTG
rialto_gene4   2279 TATGTGGATCCAGCAATGTATAGTTTCTGAACACTAGCTAGTTAACACGTTGAACTTTTG
soissons_gene4 2279 TATGTGGATCCAGCAATGTATAGTTTCTGAACACTAGCTAGTTAACACGTTGAACTTTTG
xi19_gene4     2279 TATGTGGATCCAGCAATGTATAGTTTCTGAACACTAGCTAGTTAACACGTTGAACTTTTG

cadenza_gene4  2535 CCTTTGTAGAAGAAGAGGCACAAGCACCATGCCGACGAATTGGAAGTGGTGGATCAGAAC
brompton_gene4 2535 CCTTTGTAGAAGAAGAGGCACAAGCACCATGCCGACGAATTGGAAGTGGTGGATCAGAAC
alchemy_gene4  2339 CCTTTGTAGAAGAAGAGGCACAAGCACCATGCCGACGAATTGGAAGTGGTGGATCAGAAC
rialto_gene4   2339 CCTTTGTAGAAGAAGAGGCACAAGCACCATGCCGACGAATTGGAAGTGGTGGATCAGAAC
soissons_gene4 2339 CCTTTGTAGAAGAAGAGGCACAAGCACCATGCCGACGAATTGGAAGTGGTGGATCAGAAC
xi19_gene4     2339 CCTTTGTAGAAGAAGAGGCACAAGCACCATGCCGACGAATTGGAAGTGGTGGATCAGAAC

cadenza_gene4  2595 AGAAGTTTCAATGATTTGAGAAGAGATATTGTGGTGGACCCTTGCCTAGAGCTCCGCTTC
brompton_gene4 2595 AGAAGTTTCAATGATTTGAGAAGAGATATTGTGGTGGACCCTTGCCTAGAGCTCCGCTTC
alchemy_gene4  2399 AGAAGTTTCAATGATTTGAGAAGAGATATTGTGGTGGACCCTTGCCTAGAGCTCCGCTTC
rialto_gene4   2399 AGAAGTTTCAATGATTTGAGAAGAGATATTGTGGTGGACCCTTGCCTAGAGCTCCGCTTC
soissons_gene4 2399 AGAAGTTTCAATGATTTGAGAAGAGATATTGTGGTGGACCCTTGCCTAGAGCTCCGCTTC
xi19_gene4     2399 AGAAGTTTCAATGATTTGAGAAGAGATATTGTGGTGGACCCTTGCCTAGAGCTCCGCTTC
```

```
cadenza_gene4  2655 CCCTTCCAGCCAAAGAAGGACATATCGTGCTGCCTGCAGCTCACCAACAAGGCAGCGGCC
brompton_gene4 2655 CCCTTCCAGCCAAAGAAGGACATATCGTGCTGCCTGCAGCTCACCAACAAGGCAGCGGCC
alchemy_gene4  2459 CCCTTCCAGCCAAAGAAGGACATATCGTGCTGCCTGCAGCTCACCAACAAGGCAGCGGCC
rialto_gene4   2459 CCCTTCCAGCCAAAGAAGGACATATCGTGCTGCCTGCAGCTCACCAACAAGGCAGCGGCC
soissons_gene4 2459 CCCTTCCAGCCAAAGAAGGACATATCGTGCTGCCTGCAGCTCACCAACAAGGCAGCGGCC
xi19_gene4     2459 CCCTTCCAGCCAAAGAAGGACATATCGTGCTGCCTGCAGCTCACCAACAAGGCAGCGGCC

cadenza_gene4  2715 AGCTTTGTTGCCTTCAGTATAAACACCAACGTAAACAAGTACCGCTCACTGCCAAACAAA
brompton_gene4 2715 AGCTTTGTTGCCTTCAGTATAAACACCAACGTAAACAAGTACCGCTCACTGCCAAACAAA
alchemy_gene4  2519 AGCTTTGTTGCCTTCAGTATAAACACCAACGTAAACAAGTACCGCTCACTGCCAAACAAA
rialto_gene4   2519 AGCTTTGTTGCCTTCAGTATAAACACCAACGTAAACAAGTACCGCTCACTGCCAAACAAA
soissons_gene4 2519 AGCTTTGTTGCCTTCAGTATAAACACCAACGTAAACAAGTACCGCTCACTGCCAAACAAA
xi19_gene4     2519 AGCTTTGTTGCCTTCAGTATAAACACCAACGTAAACAAGTACCGCTCACTGCCAAACAAA

cadenza_gene4  2775 GGGATTTTAGCTCCATGTTCTAAGTGCTACATCACCCTAACCTTGCGAGCGCAGCAGGAG
brompton_gene4 2775 GGGATTTTAGCTCCATGTTCTAAGTGCTACATCACCCTAACCTTGCGAGCGCAGCAGGAG
alchemy_gene4  2579 GGGATTTTAGCTCCATGTTCTAAGTGCTACATCACCCTAACCTTGCGAGCGCAGCAGGAG
rialto_gene4   2579 GGGATTTTAGCTCCATGTTCTAAGTGCTACATCACCCTAACCTTGCGAGCGCAGCAGGAG
soissons_gene4 2579 GGGATTTTAGCTCCATGTTCTAAGTGCTACATCACCCTAACCTTGCGAGCGCAGCAGGAG
xi19_gene4     2579 GGGATTTTAGCTCCATGTTCTAAGTGCTACATCACCCTAACCTTGCGAGCGCAGCAGGAG

cadenza_gene4  2835 TCACTGCCAAATATGCAGTGTGTTGATGTGTTGGTTGTGCAGGGCATAAGAGTGAGCGAG
brompton_gene4 2835 TCACTGCCAAATATGCAGTGTGTTGATGTGTTGGTTGTGCAGGGCATAAGAGTGAGCGAG
alchemy_gene4  2639 TCACTGCCAAATATGCAGTGTGTTGATGTGTTGGTTGTGCAGGGCATAAGAGTGAGCGAG
rialto_gene4   2639 TCACTGCCAAATATGCAGTGTGTTGATGTGTTGGTTGTGCAGGGCATAAGAGTGAGCGAG
soissons_gene4 2639 TCACTGCCAAATATGCAGTGTGTTGATGTGTTGGTTGTGCAGGGCATAAGAGTGAGCGAG
xi19_gene4     2639 TCACTGCCAAATATGCAGTGTGTTGATGTGTTGGTTGTGCAGGGCATAAGAGTGAGCGAG

cadenza_gene4  2895 AACTTCACGTCTGATGAAATCACTCAAGACTTCCTGGCAAATGCCGGCGCTCTGGATGAG
brompton_gene4 2895 AACTTCACGTCTGATGAAATCACTCAAGACTTCCTGGCAAATGCCGGCGCTCTGGATGAG
alchemy_gene4  2699 AACTTCACGTCTGATGAAATCACTCAAGACTTCCTGGCAAATGCCGGCGCTCTGGATGAG
rialto_gene4   2699 AACTTCACGTCTGATGAAATCACTCAAGACTTCCTGGCAAATGCCGGCGCTCTGGATGAG
soissons_gene4 2699 AACTTCACGTCTGATGAAATCACTCAAGACTTCCTGGCAAATGCCGGCGCTCTGGATGAG
xi19_gene4     2699 AACTTCACGTCTGATGAAATCACTCAAGACTTCCTGGCAAATGCCGGCGCTCTGGATGAG

cadenza_gene4  2955 GTGATGTTGCCCATTGTCTATGTTGCATTGGACCACCATCCTCTCTCTCAATAGACAGAC
brompton_gene4 2955 GTGATGTTGCCCATTGTCTATGTTGCATTGGACCACCATCCTCTCTCTCAATAGACAGAC
alchemy_gene4  2759 GTGATGTTGCCCATTGTCTATGTTGCATTGGACCACCATCCTCTCTCTCAATAGACAGAC
rialto_gene4   2759 GTGATGTTGCCCATTGTCTATGTTGCATTGGACCACCATCCTCTCTCTCAATAGACAGAC
soissons_gene4 2759 GTGATGTTGCCCATTGTCTATGTTGCATTGGACCACCATCCTCTCTCTCAATAGACAGAC
xi19_gene4     2759 GTGATGTTGCCCATTGTCTATGTTGCATTGGACCACCATCCTCTCTCTCAATAGACAGAC

cadenza_gene4  3015 AGAGGATTGTAAGAGAAACTTTTGATCTGTTCTATCAAATAGAAGGTGATATCCATAACG
brompton_gene4 3015 AGAGGATTGTAAGAGAAACTTTTGATCTGTTCTATCAAATAGAAGGTGATATCCATAACG
alchemy_gene4  2819 AGAGGATTGTAAGAGAAACTTTTGATCTGTTCTATCAAATAGAAGGTGATATCCATAACG
rialto_gene4   2819 AGAGGATTGTAAGAGAAACTTTTGATCTGTTCTATCAAATAGAAGGTGATATCCATAACG
soissons_gene4 2819 AGAGGATTGTAAGAGAAACTTTTGATCTGTTCTATCAAATAGAAGGTGATATCCATAACG
xi19_gene4     2819 AGAGGATTGTAAGAGAAACTTTTGATCTGTTCTATCAAATAGAAGGTGATATCCATAACG

cadenza_gene4  3075 AAAATCCATCTCTGCACTATCTGCGTAAGGC----------------------------
brompton_gene4 3075 AAAATCCATCTCTGCACTATCTGCGTAAGGC----------------------------
alchemy_gene4  2879 AAAATCCATCTCTGCACTATCTGCACCTGCGCTTACGGAAAAAAACCAAAATAAATACTAG
rialto_gene4   2879 AAAATCCATCTCTGCACTATCTGCACCTGCGCTTACGGAAAAAAACCAAAATAAATACTAG
soissons_gene4 2879 AAAATCCATCTCTGCACTATCTGCACCTGCGCTTACGGAAAAAAACCAAAATAAATACTAG
xi19_gene4     2879 AAAATCCATCTCTGCACTATCTGCACCTGCGCTTACGGAAAAAAACCAAAATAAATACTAG

cadenza_gene4  3106 ----ATGCTCCAAATTTCAATTTTTTTGTGGGGGTAGACAATTTGATGCGTGAGGCATGCT
brompton_gene4 3106 ----ATGCTCCAAATTTCAATTTTTTTGTGGGGGTAGACAATTTGATGCGTGAGGCATGCT
alchemy_gene4  2939 AAAAATTCAAAACATTTCAATTTTTTTGTGGGGGTAGACAATTTGATGCGTGAGGCCCGCT
rialto_gene4   2939 AAAAATTCAAAACATTTCAATTTTTTTGTGGGGGTAGACAATTTGATGCGTGAGGCCCGCT
soissons_gene4 2939 AAAAATTCAAAACATTTCAATTTTTTTGTGGGGGTAGACAATTTGATGCGTGAGGCCCGCT
xi19_gene4     2939 AAAAATTCAAAACATTTCAATTTTTTTGTGGGGGTAGACAATTTGATGCGTGAGGCCCGCT
```

```
cadenza_gene4  3162 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
brompton_gene4 3162 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
alchemy_gene4  2999 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
rialto_gene4   2999 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
soissons_gene4 2999 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
xi19_gene4     2999 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC

cadenza_gene4  3222 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
brompton_gene4 3222 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
alchemy_gene4  3059 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
rialto_gene4   3059 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
soissons_gene4 3059 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
xi19_gene4     3059 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC

cadenza_gene4  3282 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
brompton_gene4 3282 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
alchemy_gene4  3119 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
rialto_gene4   3119 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
soissons_gene4 3119 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
xi19_gene4     3119 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC

cadenza_gene4  3342 ACAATTTTTTTTTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
brompton_gene4 3342 ACAATTTTTTTTTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
alchemy_gene4  3179 ACAATTTTTT-TTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
rialto_gene4   3179 ACAATTTTTT-TTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
soissons_gene4 3179 ACAATTTTTT-TTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
xi19_gene4     3179 ACAATTTTTT-TTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA

cadenza_gene4  3402 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
brompton_gene4 3402 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
alchemy_gene4  3238 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
rialto_gene4   3238 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
soissons_gene4 3238 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
xi19_gene4     3238 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG

cadenza_gene4  3462 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
brompton_gene4 3462 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
alchemy_gene4  3298 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
rialto_gene4   3298 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
soissons_gene4 3298 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
xi19_gene4     3298 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA

cadenza_gene4  3522 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
brompton_gene4 3522 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
alchemy_gene4  3358 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
rialto_gene4   3358 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
soissons_gene4 3358 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
xi19_gene4     3358 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA

cadenza_gene4  3582 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
brompton_gene4 3582 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
alchemy_gene4  3418 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
rialto_gene4   3418 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
soissons_gene4 3418 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
xi19_gene4     3418 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA

cadenza_gene4  3642 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCATGCT
brompton_gene4 3642 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCATGCT
alchemy_gene4  3478 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCCCGCT
rialto_gene4   3478 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCCCGCT
soissons_gene4 3478 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCCCGCT
xi19_gene4     3478 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCCCGCT
```

```
cadenza_gene4  3162 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
brompton_gene4 3162 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
alchemy_gene4  2999 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
rialto_gene4   2999 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
soissons_gene4 2999 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC
xi19_gene4     2999 CCAAATTTCAAGTCATTTGGACGTTTGATCAGCTCTCAGCAAAAAGACAAATTGGGGGGTC

cadenza_gene4  3222 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
brompton_gene4 3222 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
alchemy_gene4  3059 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
rialto_gene4   3059 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
soissons_gene4 3059 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC
xi19_gene4     3059 TGTAAAAATGTTCACTGTTCAGGTACTGTTTTGAACCGATTTGTCTTTTTTGCTGAGAGC

cadenza_gene4  3282 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
brompton_gene4 3282 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
alchemy_gene4  3119 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
rialto_gene4   3119 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
soissons_gene4 3119 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC
xi19_gene4     3119 TGCTCAAATGTCCAAATGACTTGAAATTTGGAGCGGGCCTCACACATGAAATTGTCTACC

cadenza_gene4  3342 ACAATTTTTTTTTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
brompton_gene4 3342 ACAATTTTTTTTTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
alchemy_gene4  3179 ACAATTTTTT-TTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
rialto_gene4   3179 ACAATTTTTT-TTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
soissons_gene4 3179 ACAATTTTTT-TTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA
xi19_gene4     3179 ACAATTTTTT-TTTGGAATTTGTTGAATTTGTTTTGAATCTTTTACGAATTTTTTTCTAA

cadenza_gene4  3402 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
brompton_gene4 3402 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
alchemy_gene4  3238 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
rialto_gene4   3238 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
soissons_gene4 3238 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG
xi19_gene4     3238 CCGGGTGCAGATGAGCCTGGGCACCGAAACGCCCTACTCCATAACGAAAATCCATCTCTG

cadenza_gene4  3462 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
brompton_gene4 3462 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
alchemy_gene4  3298 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
rialto_gene4   3298 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
soissons_gene4 3298 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA
xi19_gene4     3298 CACTAGGACTATGTACGGCTTGGACGTTGATCTTCCTGAAATCAAAAAAATTCTTCACCA

cadenza_gene4  3522 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
brompton_gene4 3522 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
alchemy_gene4  3358 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
rialto_gene4   3358 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
soissons_gene4 3358 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA
xi19_gene4     3358 ATATGCATGTTCAACTTCACAACTTTGATAAAGTTTAAGCACCATTCTGTTTTGTTAAAA

cadenza_gene4  3582 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
brompton_gene4 3582 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
alchemy_gene4  3418 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
rialto_gene4   3418 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
soissons_gene4 3418 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA
xi19_gene4     3418 CCAGTCTCAACATTGAATCCTCAGGTGCATGTATGTTCCAGACAAAATGAACCGATGAAA

cadenza_gene4  3642 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCATGCT
brompton_gene4 3642 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCATGCT
alchemy_gene4  3478 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCCCGCT
rialto_gene4   3478 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCCCGCT
soissons_gene4 3478 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCCCGCT
xi19_gene4     3478 AATGTTTCCAGCATATTTTGTTTTGCAACATGTCAAATTGTACACGAACAAGGGCCCGCT
```

C

Chapter 4 Supplementary Material

Table C.1 – Phenotypes observed in the Avalon × Cadenza field trials

Line	2016						2017				
	Height	**HD (dap)**	**1° Stago**	**2° Stago**	**3° Stago**	**Glume**	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**Glume**
AxC 001	73.0	66.0	25.0	30.0	75.0	50.0	75.0	258.0	52.5	92.5	7.5
AxC 002	47.5	68.5	35.0	42.5	96.5	99.0	48.0	261.0	25.0	75.0	20.0
AxC 003	-	-	-	-	-	-	63.5	266.0	20.0	70.0	15.0
AxC 005	61.0	-	15.0	15.0	50.0	10.0	45.0	260.5	10.0	50.0	0.0
AxC 006	68.0	65.5	25.0	42.5	98.0	72.5	-	262.5	-	-	-
AxC 007	76.0	64.0	22.5	27.5	80.0	17.5	60.0	258.0	25.0	80.0	15.0
AxC 008	64.0	59.5	30.0	30.0	72.5	62.5	60.0	262.0	32.5	67.5	22.5
AxC 009	-	-	-	-	-	-	59.0	260.5	40.0	75.0	5.0
AxC 010	-	-	-	-	-	-	64.5	260.0	30.0	77.5	3.5
AxC 011	-	-	-	-	-	-	47.5	256.0	70.0	67.5	15.0
AxC 012	-	-	-	-	-	-	69.0	259.0	32.5	75.0	7.5
AxC 013	-	-	-	-	-	-	69.0	260.5	30.0	75.0	4.0
AxC 014	-	-	-	-	-	-	48.0	259.5	47.5	65.0	10.0
AxC 015	-	-	-	-	-	-	64.0	261.0	37.5	75.0	20.0
AxC 016	61.0	66.5	35.0	37.5	96.5	32.5	50.5	265.0	27.5	57.5	11.5
AxC 017	57.0	66.5	22.5	45.0	94.0	92.5	40.0	265.5	-	75.0	-
AxC 018	77.5	64.0	27.5	30.0	90.0	55.0	67.0	262.5	22.5	55.0	12.5
AxC 019	80.0	68.0	12.5	15.0	40.0	12.5	79.0	261.0	20.0	80.0	7.5
AxC 020	-	-	-	-	-	-	64.0	260.0	37.5	52.5	8.5
AxC 021	61.0	67.0	32.5	40.0	99.0	60.0	64.5	262.0	47.5	80.0	3.5
AxC 022	61.0	66.5	35.0	40.0	95.0	82.5	60.0	260.5	42.5	80.0	7.5
AxC 023	80.5	62.0	17.5	17.5	55.0	17.5	61.0	257.5	12.5	57.5	3.5
AxC 024	55.5	64.0	22.5	25.0	75.0	75.0	54.5	256.0	37.5	82.5	12.5

Continued on next page

Table C.1 – continued from previous page

	2016						2017				
Line	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**3° Stago**	**Glume**	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**Glume**
AxC 025	-	-	-	-	-	-	52.0	261.0	55.0	82.5	10.0
AxC 026	59.0	-	10.0	15.0	75.0	70.0	59.5	262.5	37.5	67.5	17.5
AxC 027	50.0	69.0	25.0	27.5	72.5	65.0	49.0	265.0	20.0	40.0	20.0
AxC 028	60.0	69.0	12.5	20.0	52.5	50.0	67.0	265.0	32.5	67.5	15.0
AxC 029	-	-	-	-	-	-	53.5	262.5	37.5	77.5	22.5
AxC 030	-	-	-	-	-	-	54.0	263.5	45.0	82.5	12.5
AxC 031	-	-	-	-	-	-	58.0	265.0	15.0	50.0	10.0
AxC 032	-	-	-	-	-	-	47.5	258.0	20.0	60.0	7.5
AxC 033	-	-	-	-	-	-	68.0	260.5	45.0	80.0	20.0
AxC 034	-	-	-	-	-	-	84.5	257.5	75.0	95.0	2.5
AxC 035	-	-	-	-	-	-	44.0	261.0	55.0	90.0	20.0
AxC 036	-	-	-	-	-	-	70.0	260.5	35.0	57.5	4.0
AxC 037	-	-	-	-	-	-	73.0	257.0	35.0	80.0	7.5
AxC 038	-	-	-	-	-	-	68.5	260.5	37.5	82.5	5.0
AxC 039	-	-	-	-	-	-	64.0	263.5	27.5	65.0	10.0
AxC 041	80.0	62.5	15.0	25.0	62.5	12.5	72.0	257.5	25.0	72.5	7.5
AxC 042	70.0	69.0	20.0	25.0	85.0	45.0	62.0	263.5	42.5	70.0	10.0
AxC 043	71.5	66.0	17.5	20.0	72.5	32.5	56.0	268.0	10.0	30.0	5.0
AxC 044	71.5	69.0	32.5	32.5	82.5	6.0	68.0	266.0	35.0	60.0	3.5
AxC 046	65.0	64.0	25.0	30.0	89.0	75.0	64.0	259.0	45.0	87.5	9.0
AxC 047	-	-	-	-	-	-	48.0	264.5	15.0	30.0	10.0
AxC 049	-	-	-	-	-	-	58.0	260.0	47.5	90.0	14.0
AxC 050	-	-	-	-	-	-	66.0	259.5	42.5	67.5	10.0

Continued on next page

Table C.1 – continued from previous page

	2016						2017				
Line	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**3° Stago**	**Glume**	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**Glume**
AxC 051	-	-	-	20.0	-	-	63.5	259.0	22.5	55.0	17.5
AxC 052	55.0	67.5	30.0	35.0	90.0	22.5	51.0	263.5	30.0	60.0	2.0
AxC 053	61.5	64.0	22.5	30.0	72.5	27.5	60.0	262.0	20.0	60.0	8.0
AxC 054	59.3	63.5	17.5	25.0	87.5	62.5	53.5	264.0	27.5	47.5	12.5
AxC 055	77.0	68.0	12.5	20.0	62.5	22.5	74.5	258.5	30.0	65.0	9.5
AxC 056	72.0	62.0	15.0	32.5	82.5	15.0	67.5	259.0	25.0	50.0	6.5
AxC 057	-	-	-	-	-	-	59.5	266.5	30.0	70.0	10.0
AxC 058	68.0	65.0	22.5	30.0	87.5	22.5	64.0	263.0	37.5	65.0	9.0
AxC 059	-	-	-	-	-	-	57.0	258.5	50.0	77.5	11.5
AxC 060	-	-	-	-	-	-	61.5	258.5	80.0	97.5	5.0
AxC 061	-	-	-	-	-	-	58.5	260.0	40.0	80.0	12.5
AxC 062	83.0	63.5	27.5	30.0	72.5	12.5	82.0	260.0	32.5	70.0	2.5
AxC 063	-	-	-	-	-	-	53.0	267.5	22.5	62.5	10.0
AxC 064	-	-	-	-	-	-	59.5	261.5	55.0	87.5	7.5
AxC 065	69.5	62.0	27.5	35.0	87.5	57.5	65.0	259.0	30.0	60.0	11.0
AxC 066	-	-	-	-	-	-	52.0	265.5	35.0	82.5	20.0
AxC 067	-	-	-	60.0	-	-	74.0	260.0	37.5	72.5	5.0
AxC 068	-	-	-	-	-	-	58.0	261.0	40.0	80.0	5.0
AxC 069	-	-	-	-	-	-	75.5	260.5	50.0	82.5	11.5
AxC 070	58.5	65.0	17.5	22.5	75.0	60.0	60.0	260.5	25.0	60.0	7.5
AxC 071	86.0	69.0	17.5	17.5	47.5	17.5	86.5	260.5	30.0	75.0	10.0
AxC 072	-	-	-	-	-	-	61.0	260.5	42.5	77.5	9.0
AxC 073	83.8	61.0	22.5	27.5	87.5	40.0	78.0	258.5	40.0	85.0	3.5

Continued on next page

Table C.1 – continued from previous page

	2016						2017				
Line	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**3° Stago**	**Glume**	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**Glume**
AxC 074	48.5	68.5	35.0	47.5	79.0	69.0	42.5	264.5	15.0	37.5	20.0
AxC 075	-	-	-	-	-	-	59.0	258.5	55.0	80.0	15.0
AxC 076	-	-	-	-	-	-	85.0	260.5	35.0	65.0	7.5
AxC 077	-	-	-	-	-	-	58.5	261.0	25.0	67.5	9.0
AxC 078	-	-	-	-	-	-	62.0	261.0	42.5	82.5	10.0
AxC 079	85.5	65.0	17.5	20.0	65.0	20.0	78.0	258.5	27.5	80.0	9.0
AxC 080	49.5	66.5	25.0	30.0	89.0	90.0	47.0	260.0	57.5	87.5	27.5
AxC 081	70.5	66.5	25.0	27.5	80.0	35.0	69.0	260.5	37.5	67.5	6.0
AxC 083	-	-	-	-	-	-	79.0	260.0	35.0	70.0	2.5
AxC 084	78.5	64.0	20.0	30.0	87.5	47.5	72.0	260.0	15.0	80.0	5.0
AxC 085	74.5	67.5	27.5	35.0	77.5	5.0	68.0	266.0	30.0	57.5	5.0
AxC 086	-	-	-	-	-	-	66.5	260.5	57.5	90.0	10.0
AxC 087	65.5	66.5	17.5	22.5	67.5	45.0	55.5	260.5	22.5	32.5	10.0
AxC 088	59.0	66.0	22.5	27.5	50.0	72.5	59.0	258.5	75.0	90.0	10.0
AxC 089	-	-	-	-	-	-	51.5	258.5	35.0	70.0	11.5
AxC 090	-	-	-	-	-	-	67.0	259.5	37.5	77.5	6.5
AxC 091	-	-	-	-	-	-	52.0	260.5	52.5	85.0	20.0
AxC 092	53.0	65.0	35.0	50.0	91.5	50.0	57.0	261.0	25.0	67.5	13.5
AxC 093	62.0	62.0	32.5	82.5	100.0	87.5	57.0	257.0	57.5	85.0	20.0
AxC 094	-	-	-	-	-	-	58.0	263.5	37.5	80.0	10.0
AxC 096	62.5	65.0	35.0	37.5	90.0	60.0	64.0	263.5	20.0	67.5	11.5
AxC 097	71.5	63.0	37.5	40.0	90.0	25.0	71.5	260.5	37.5	80.0	3.5
AxC 098	-	-	-	-	-	-	61.5	263.0	47.5	77.5	7.5

Continued on next page

Table C.1 – continued from previous page

	2016						2017				
Line	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**3° Stago**	**Glume**	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**Glume**
AxC 099	75.5	66.0	25.0	32.5	87.5	30.0	71.0	258.5	62.5	80.0	10.0
AxC 100	-	-	-	-	-	-	77.0	261.0	52.5	77.5	8.0
AxC 101	56.5	68.0	27.5	37.5	97.5	95.0	64.0	262.0	50.0	85.0	15.0
AxC 102	66.5	63.0	32.5	47.5	98.0	94.0	60.5	260.0	25.0	75.0	25.0
AxC 103	55.5	66.0	45.0	55.0	96.5	65.0	57.0	261.0	47.5	90.0	10.0
AxC 104	71.5	68.0	25.0	27.5	60.0	20.0	58.5	261.0	30.0	50.0	5.0
AxC 105	65.5	66.5	20.0	25.0	75.0	52.5	57.0	266.0	5.0	40.0	10.0
AxC 106	-	-	-	-	-	-	55.5	260.5	30.0	65.0	10.0
AxC 107	71.5	64.5	25.0	30.0	91.5	20.0	69.5	259.0	52.5	95.0	5.0
AxC 108	63.3	63.0	25.0	37.5	87.5	85.0	54.5	261.5	37.5	75.0	25.0
AxC 109	51.0	70.0	25.0	32.5	70.0	52.5	-	-	-	-	-
AxC 110	-	-	-	-	-	-	72.0	260.0	35.0	77.5	7.5
AxC 111	-	-	-	-	-	-	59.0	260.5	52.5	85.0	10.0
AxC 112	-	-	-	15.0	-	-	59.0	263.0	32.5	67.5	6.5
AxC 113	-	-	-	-	-	-	52.5	259.5	45.0	82.5	7.5
AxC 114	52.5	65.0	20.0	37.5	87.5	72.5	57.5	258.0	57.5	90.0	9.0
AxC 115	80.0	66.0	30.0	30.0	95.0	40.0	73.0	261.0	40.0	80.0	10.0
AxC 116	47.0	69.0	35.0	40.0	87.5	75.0	45.5	268.0	22.5	67.5	25.0
AxC 117	60.0	67.5	40.0	52.5	99.0	42.5	46.5	260.0	37.5	65.0	20.0
AxC 118	-	-	-	-	-	-	65.5	260.0	50.0	82.5	20.0
AxC 119	63.3	63.5	20.0	25.0	89.0	65.0	65.0	257.5	67.5	87.5	5.0
AxC 120	67.5	67.0	25.0	27.5	80.0	50.0	45.0	260.0	10.0	70.0	45.0
AxC 121	-	-	-	-	-	-	68.5	261.0	45.0	87.5	15.0

Continued on next page

Table C.1 – continued from previous page

	2016						2017				
Line	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**3° Stago**	**Glume**	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**Glume**
AxC 122	67.0	59.0	25.0	30.0	90.0	20.0	66.5	266.0	32.5	75.0	6.5
AxC 123	-	-	-	-	-	-	72.5	258.0	22.5	62.5	7.5
AxC 124	-	-	-	-	-	-	63.5	265.0	32.5	70.0	7.5
AxC 125	61.0	68.0	20.0	20.0	72.5	72.5	63.5	261.0	17.5	57.5	7.5
AxC 126	68.0	62.5	22.5	32.5	95.0	20.0	50.0	260.0	20.0	50.0	5.0
AxC 127	-	-	-	-	-	-	62.5	264.5	42.5	60.0	12.5
AxC 128	-	-	-	-	-	-	62.5	260.5	25.0	72.5	7.5
AxC 129	66.0	67.5	25.0	30.0	65.0	22.5	69.0	262.0	20.0	67.5	10.0
AxC 130	53.0	68.5	17.5	27.5	82.5	62.5	47.0	266.5	15.0	82.5	25.0
AxC 131	-	-	-	-	-	-	58.5	261.0	67.5	87.5	12.5
AxC 132	51.5	60.5	47.5	47.5	94.0	92.5	-	-	20.0	-	-
AxC 133	70.5	66.5	30.0	35.0	75.0	25.0	56.0	265.0	25.0	50.0	10.0
AxC 134	54.5	66.5	22.5	37.5	98.0	92.5	47.5	258.5	27.5	92.5	20.0
AxC 136	-	-	-	-	-	-	56.0	258.5	67.5	85.0	16.0
AxC 137	79.3	67.5	22.5	25.0	80.0	17.5	67.5	260.0	45.0	65.0	12.5
AxC 138	59.0	65.0	15.0	25.0	90.0	87.5	60.5	260.0	35.0	57.5	20.0
AxC 139	65.5	65.0	27.5	30.0	94.0	35.0	65.0	260.0	52.5	87.5	6.5
AxC 140	79.0	61.0	22.5	25.0	92.5	67.5	70.0	256.0	32.5	65.0	6.5
AxC 141	67.5	67.5	20.0	27.5	50.0	45.0	52.0	262.0	20.0	60.0	12.5
AxC 142	-	-	-	-	-	-	74.0	264.0	57.5	82.5	3.5
AxC 143	58.5	68.5	15.0	22.5	62.5	60.0	54.5	263.0	17.5	55.0	10.0
AxC 144	-	-	-	-	-	-	68.0	258.0	42.5	85.0	7.5
AxC 145	70.5	66.0	22.5	32.5	95.0	55.0	58.0	260.0	40.0	32.5	15.0

Continued on next page

Table C.1 – continued from previous page

	2016						2017				
Line	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**3° Stago**	**Glume**	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**Glume**
AxC 146	54.5	68.0	27.5	32.5	90.0	52.5	46.5	263.5	42.5	62.5	12.5
AxC 147	57.0	68.5	37.5	55.0	97.5	90.0	46.0	267.0	25.0	40.0	40.0
AxC 148	68.0	62.5	25.0	35.0	95.0	47.5	65.0	256.5	35.0	70.0	10.0
AxC 149	-	-	-	-	-	-	72.5	258.0	65.0	92.5	10.0
AxC 150	80.0	69.5	17.5	20.0	52.5	20.0	76.0	263.0	50.0	80.0	10.0
AxC 151	-	-	-	-	-	-	47.0	263.5	50.0	95.0	5.0
AxC 152	60.0	68.0	35.0	40.0	80.0	95.0	61.5	261.0	40.0	70.0	12.5
AxC 153	-	-	-	-	-	-	56.0	260.0	47.5	87.5	15.0
AxC 154	-	-	-	10.0	-	-	48.5	257.0	65.0	90.0	22.5
AxC 155	-	-	-	-	-	-	65.5	259.0	47.5	90.0	17.5
AxC 156	64.0	68.5	17.5	22.5	55.0	60.0	-	-	-	-	-
AxC 157	-	-	-	-	-	-	76.0	254.5	57.5	91.5	7.5
AxC 158	57.0	68.5	22.5	37.5	77.5	60.0	54.5	261.0	30.0	60.0	22.5
AxC 159	55.0	68.5	32.5	35.0	92.5	82.5	60.0	261.0	42.5	77.5	11.0
AxC 160	-	-	-	-	-	-	58.5	261.0	42.5	82.5	7.5
AxC 161	-	-	-	-	-	-	62.5	260.0	35.0	77.5	10.0
AxC 162	70.5	-	45.0	40.0	96.5	35.0	71.5	258.0	55.0	80.0	16.5
AxC 163	53.0	68.5	25.0	27.5	90.0	65.0	53.5	266.0	27.5	55.0	12.5
AxC 164	69.0	67.5	30.0	35.0	95.0	87.5	68.5	260.5	47.5	85.0	12.5
AxC 165	71.5	65.0	17.5	30.0	95.0	25.0	62.5	260.0	30.0	75.0	7.5
AxC 166	73.5	61.0	27.5	60.0	99.0	67.5	79.0	256.5	67.5	90.0	5.0
AxC 167	54.0	66.0	15.0	20.0	50.0	85.0	57.0	262.0	22.5	50.0	27.5
AxC 168	65.5	67.5	25.0	35.0	85.0	40.0	65.0	263.0	35.0	80.0	8.0

Continued on next page

Table C.1 – continued from previous page

	2016						2017				
Line	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**3° Stago**	**Glume**	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**Glume**
AxC 169	-	-	-	-	-	-	63.5	262.0	22.5	62.5	12.5
AxC 170	62.5	64.5	25.0	37.5	96.5	92.5	51.0	265.0	12.5	40.0	12.5
AxC 171	-	-	-	-	-	-	66.0	256.0	55.0	90.0	5.0
AxC 172	64.5	63.5	22.5	27.5	74.0	12.5	65.5	260.0	15.0	47.5	10.0
AxC 173	61.5	67.5	27.5	35.0	92.5	62.5	50.0	269.0	15.0	60.0	5.0
AxC 174	-	-	-	-	-	-	55.0	262.0	37.5	67.5	20.0
AxC 175	-	-	-	-	-	-	53.5	259.5	65.0	85.0	10.0
AxC 176	55.0	66.0	22.5	32.5	82.5	67.5	51.5	264.5	12.5	37.5	17.5
AxC 177	-	-	-	-	-	-	70.5	260.0	40.0	82.5	5.0
AxC 178	-	-	-	-	-	-	64.5	260.5	30.0	72.5	14.0
AxC 179	-	-	-	-	-	-	44.5	259.0	37.5	80.0	20.0
AxC 180	68.5	66.0	15.0	20.0	37.5	25.0	69.5	262.0	37.5	75.0	2.5
AxC 181	-	-	-	-	-	-	70.0	260.5	37.5	77.5	7.5
AxC 182	-	-	-	-	-	-	67.0	261.0	37.5	75.0	15.0
AxC 183	70.0	68.0	10.0	13.5	37.5	35.0	62.0	264.0	35.0	77.5	12.5
AxC 184	-	-	20.0	-	-	15.0	52.5	264.0	37.5	22.5	10.0
AxC 185	-	-	-	-	-	-	67.0	261.0	42.5	80.0	15.0
AxC 186	60.8	63.5	22.5	52.5	95.0	55.0	57.0	258.5	40.0	75.0	20.0
AxC 187	-	-	-	-	-	-	69.0	256.5	75.0	92.5	10.0
AxC 188	73.5	70.0	20.0	20.0	52.5	25.0	68.0	263.5	32.5	72.5	12.5
AxC 189	-	-	-	-	-	-	84.0	260.0	40.0	90.0	5.0
AxC 190	-	-	-	-	-	-	84.5	262.0	70.0	85.0	4.0
AxC 191	84.0	66.0	15.0	20.0	80.0	17.5	72.0	266.0	22.5	27.5	11.0

Continued on next page

Table C.1 – continued from previous page

	2016						2017				
Line	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**3° Stago**	**Glume**	**Height (cm)**	**HD (dap)**	**1° Stago**	**2° Stago**	**Glume**
AxC 192	52.0	67.0	27.5	60.0	99.0	90.0	43.0	264.0	20.0	25.0	5.0
AxC 193	52.8	68.5	15.0	25.0	72.5	52.5	46.5	266.0	42.5	77.5	22.5
AxC 194	58.0	67.5	32.5	35.0	91.5	55.0	47.5	262.0	10.0	25.0	13.5
AxC 195	82.3	64.0	15.0	20.0	60.0	5.0	71.0	260.0	30.0	65.0	3.5
AxC 196	56.5	65.5	15.0	22.5	85.0	45.0	59.0	260.0	20.0	30.0	5.0
AxC 197	-	-	-	-	-	-	69.0	260.0	65.0	87.5	7.5
AxC 198	-	-	-	-	-	-	67.0	260.0	47.5	85.0	15.0
AxC 199	55.5	67.0	15.0	22.5	87.5	65.0	51.5	263.5	7.5	30.0	8.0
AxC 200	79.0	62.0	15.0	20.0	77.5	25.0	69.0	260.0	50.0	85.0	20.0
AxC 202	66.0	62.5	25.0	50.0	95.0	50.0	60.0	258.5	35.0	57.5	10.0
AxC 203	81.5	67.0	22.5	22.5	55.0	10.0	80.5	259.0	12.5	65.0	6.5
AxC 204	-	-	-	-	-	-	54.5	265.0	47.5	75.0	14.0
AxC 205	63.5	65.5	17.5	22.5	70.0	72.5	59.5	260.5	30.0	60.0	4.0
AxC 206	-	-	-	-	-	-	74.5	257.5	47.5	90.0	6.5

Field trial average phenotype data for 198 lines of the bi-parental mapping population Avalon × Cadenza in 2016 and 2017. HD = heading date. dap = days after planting. Stago = percentage coverage of leaf blotch on the plot. Glume = percentage coverage of glume blotch on the plot. Primary = first timepoint. Secondary = second timepoint. Teritary = third timepoint.

D

Additional Supplementary Material

Table D.1 – g187 and g188 allele calls for a small panel of Australian commercial lines

Line	SnTox3 Score	S.E.	g188-1	g188-2	g187-1	g187-2
Line 13	0	0	Allele 1	Allele 1	Allele 1	Allele 1
Line 55	2.6	0.24	Allele 1	Heterozygote	Allele 2	Allele 2
Line 27	3.2	0.8	Allele 1	Heterozygote	Allele 2	Allele 2
Line 34	3.2	0.8	Allele 1	Heterozygote	No Call	Heterozygote
Line 23	3.6	0.24	Allele 1	Heterozygote	Allele 2	Heterozygote
Line 68	3.6	0.24	Allele 1	Heterozygote	Allele 2	Allele 2
Line 4	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 16	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 18	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 21	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 24	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 25	4	0	Allele 1	Heterozygote	Heterozygote	Allele 1
Line 32	4	0	Allele 1	Allele 1	Heterozygote	Allele 1
Line 40	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 41	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 42	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 45	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 47	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 48	4	0	Allele 1	Heterozygote	Heterozygote	Allele 2
Line 50	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 58	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 63	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 64	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 67	4	0	Allele 1	Heterozygote	No Call	No Call
Line 70	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 71	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 74	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 83	4	0	Allele 1	Heterozygote	Allele 2	Allele 2
Line 94	4	0	Allele 1	Heterozygote	Allele 2	Allele 1
Line 2	0	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 10	1.6	0.98	Allele 2	Allele 1	No Call	Allele 1
Line 72	2.4	0.98	Allele 2	Allele 1	Heterozygote	Allele 1
Line 1	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 3	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 5	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 6	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 11	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 12	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 14	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 15	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 17	4	0	Allele 2	Allele 1	Allele 1	Allele 1

Continued on next page

Table D.1 – continued from previous page

Line	SnTox3 Score	S.E.	g188-1	g188-2	g187-1	g187-2
Line 19	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 22	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 26	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 28	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 29	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 31	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 33	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 35	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 36	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 38	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 39	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 43	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 49	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 51	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 52	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 53	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 54	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 57	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 61	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 62	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 65	4	0	Allele 2	Allele 1	Allele 1	Heterozygote
Line 66	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 73	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 75	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 76	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 77	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 78	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 79	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 80	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 82	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 85	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 87	4	0	Allele 2	Allele 1	Allele 1	No Call
Line 88	4	0	Allele 2	Allele 1	Undetermined	Allele 1
Line 89	4	0	Allele 2	Allele 1	Allele 1	Allele 1
Line 90	4	0	Allele 2	Allele 1	Undetermined	No Call
Line 92	4	0	Allele 2	Allele 1	No Call	Allele 1
Line 7	0	0	No Call	Allele 1	Allele 1	Allele 1
Line 8	0	0	No Call	Allele 1	Allele 1	Allele 1
Line 20	0	0	No Call	Allele 1	Allele 1	Allele 1
Line 37	0	0	No Call	Allele 1	Allele 2	Allele 1
Line 44	0	0	No Call	Allele 1	Allele 1	Allele 1
Line 46	0	0	No Call	Allele 1	No Call	Allele 1

Continued on next page

Table D.1 – continued from previous page

Line	SnTox3 Score	S.E.	g188-1	g188-2	g187-1	g187-2
Line 56	0	0	No Call	Allele 1	Allele 1	Allele 1
Line 59	0	0	No Call	Allele 1	Allele 2	No Call
Line 60	0	0	No Call	Allele 1	Allele 1	Allele 1
Line 69	0	0	No Call	Allele 1	Allele 1	Allele 1
Line 86	0	0	No Call	Allele 1	Allele 1	Allele 1
Line 93	0	0	No Call	Allele 1	Allele 1	Allele 1
Line 9	0.8	0.8	No Call	Allele 1	Allele 2	Allele 1
Line 91	4	0	No Call	Allele 1	Allele 2	No Call
Line 81	-	-	Allele 1	Heterozygote	Allele 2	No Call
Line 84	-	-	Allele 2	Allele 1	Allele 1	Allele 1
Line 30	4	0	No Call	No Call	No Call	No Call

The KASP calls for two KASP markers designed for allele calling of g188 (g188-1 and g188-2) and g187 (g187-1 and g187-2). g188-1 was designed with allele 1 based on the intermediately sensitive version of g188 and allele 2 based on the highly sensitive version of g188. In the majority of cases, with the KASP marker g188-1 allele 1 and allele 2 are associated with lines highly sensitive to SnTox3, but No Call is associated with insensitivity to SnTox3. This analysis was carried out by collaborators at Curtin University.

E

Published Literature

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 July 2018
doi: 10.3389/fpls.2018.00881

Assessing European Wheat Sensitivities to *Parastagonospora nodorum* Necrotrophic Effectors and Fine-Mapping the *Snn3-B1* Locus Conferring Sensitivity to the Effector SnTox3

Rowena C. Downie[1,2]*, Laura Bouvet[1,2], Eiko Furuki[3], Nick Gosman[1†], Keith A. Gardner[1], Ian J. Mackay[1], Camila Campos Mantello[1], Greg Mellers[1], Huyen T. T. Phan[3], Gemma A. Rose[1], Kar-Chun Tan[3], Richard P. Oliver[3] and James Cockram[1]*

[1] Genetics and Breeding Department, National Institute of Agricultural Botany, Cambridge, United Kingdom, [2] Plant Sciences Department, University of Cambridge, Cambridge, United Kingdom, [3] Centre for Crop and Disease Management, Curtin University, Perth, WA, Australia

OPEN ACCESS

Edited by:
Luigi Cattivelli, Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia Agraria (CREA), Italy

Reviewed by:
Agata Gadaleta, Università degli studi di Bari Aldo Moro, Italy
Alessandro Tondelli, Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia Agraria (CREA), Italy

***Correspondence:**
Rowena C. Downie
rowena.downie@niab.com
James Cockram
james.cockram@niab.com

†Present address:
Nick Gosman, Azotic Technologies Ltd., Nottingham, United Kingdom

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *22 January 2018*
Accepted: *06 June 2018*
Published: *04 July 2018*

Citation:
Downie RC, Bouvet L, Furuki E, Gosman N, Gardner KA, Mackay IJ, Campos Mantello C, Mellers G, Phan HTT, Rose GA, Tan K-C, Oliver RP and Cockram J (2018) Assessing European Wheat Sensitivities to Parastagonospora nodorum Necrotrophic Effectors and Fine-Mapping the Snn3-B1 Locus Conferring Sensitivity to the Effector SnTox3. Front. Plant Sci. 9:881. doi: 10.3389/fpls.2018.00881

Parastagonospora nodorum is a necrotrophic fungal pathogen of wheat (*Triticum aestivum* L.), one of the world's most important crops. *P. nodorum* mediates host cell death using proteinaceous necrotrophic effectors, presumably liberating nutrients that allow the infection process to continue. The identification of pathogen effectors has allowed host genetic resistance mechanisms to be separated into their constituent parts. In *P. nodorum*, three proteinaceous effectors have been cloned: *SnToxA*, *SnTox1*, and *SnTox3*. Here, we survey sensitivity to all three effectors in a panel of 480 European wheat varieties, and fine-map the wheat SnTox3 sensitivity locus *Snn3-B1* using genome-wide association scans (GWAS) and an eight-founder wheat multi-parent advanced generation inter-cross (MAGIC) population. Using a Bonferroni corrected $P \leq 0.05$ significance threshold, GWAS identified 10 significant markers defining a single locus, *Snn3-B1*, located on the short arm of chromosome 5B explaining 32% of the phenotypic variation [peak single nucleotide polymorphisms (SNPs), Excalibur_c47452_183 and GENE-3324_338, $-\log_{10}P = 20.44$]. Single marker analysis of SnTox3 sensitivity in the MAGIC population located *Snn3-B1* via five significant SNPs, defining a 6.2-kb region that included the two peak SNPs identified in the association mapping panel. Accordingly, SNP Excalibur_c47452_183 was converted to the KASP genotyping system, and validated by screening a subset of 95 wheat varieties, providing a valuable resource for marker assisted breeding and for further genetic investigation. In addition, composite interval mapping in the MAGIC population identified six minor SnTox3 sensitivity quantitative trait loci, on chromosomes 2A (*QTox3.niab-2A.1*, *P*-value = 9.17^{-7}), 2B (*QTox3.niab-2B.1*, $P = 0.018$), 3B (*QTox3.niab-3B.1*, $P = 48.51^{-4}$), 4D (*QTox3.niab-4D.1*, $P = 0.028$), 6A (*QTox3.niab-6A.1*, $P = 8.51^{-4}$), and 7B (*QTox3.niab-7B.1*, $P = 0.020$), each accounting for between 3.1 and 6.0 % of

the phenotypic variance. Collectively, the outcomes of this study provides breeders with knowledge and resources regarding the sensitivity of European wheat germplasm to *P. nodorum* effectors, as well as simple diagnostic markers for determining allelic state at *Snn3-B1*.

Keywords: *Parastagonospora nodorum*, SnTox3, fungal effector proteins, multi-parent advanced generation inter-cross, genome-wide association scans, high-density SNP genotyping, genetic markers

INTRODUCTION

The necrotrophic pathogen *Parastagonospora* (synonyms: *Septoria*, *Stagonospora*, *Phaeosphaeria*) *nodorum* (Berk.) Quaedvlieg, Verkley, and Crous is the causal agent of the disease SNB and glume blotch in wheat (*Triticum aestivum* L.), a disease of significant economic importance in Australia, Europe, North America, and Northern Africa (Friesen et al., 2005; Oliver et al., 2012; Quaedvlieg et al., 2013). The visual symptoms of SNB are chlorosis and necrosis of the leaf tissue, as well as discoloration of the glumes, often in the form of lesions (Solomon et al., 2006). These symptoms reduce the leaf surface area capable of photosynthesis, limiting overall crop growth, with SNB shown to result in grain yield losses of up to 31% (Bhathal and Loughman, 2003). *P. nodorum* is thought to derive nutrients from dead plant tissue, utilizing fungal effector proteins, previously known as host-specific (or selective) toxins, to induce a hypersensitive response in the host, which takes the form of programmed cell death (Friesen et al., 2007; Liu et al., 2009; Oliver et al., 2012). The necrotic response in the sensitive host is hypothesized to facilitate pathogen colonization, promoting infection and ultimately providing a rich nutrient source, via cell death (Oliver and Solomon, 2010; Vincent et al., 2012). This is known as effector-triggered susceptibility and is genetically induced via an "inverse gene for gene system" (Friesen et al., 2007). The identification of effector proteins has led to a paradigm shift in the approach to tackle these types of pathogens, as the host–pathogen interactions can be broken down into their constituent parts. Consequently, targeted breeding could then be used to eliminate host sensitivity on an effector by effector basis.

Effector proteins were described for the first time with regards to a host–pathogen interaction between *Alternaria alternata* (a necrotroph) and *Pirus serotine* (Tanaka, 1933). However, the first protein effector described in a necrotrophic pathogen was PtrToxA from the wheat tan spot pathogen, *Pyrenophora tritici-repentis*, which triggers necrosis in wheat lines carrying susceptible alleles at the *Tsn1* locus (Ballance et al., 1989; Tomas et al., 1990; Faris et al., 1996). A near identical effector, SnToxA, was discovered in *P. nodorum*, with the corresponding host sensitivity locus also being *Tsn1* (Liu et al., 2006). *Tsn1*, located on the long arm of chromosome 5B, has been cloned and encodes an intracellular protein with a serine/threonine protein kinase (S/TPK) domain, a nucleotide-binding site (NBS), and leucine-rich repeats (LRRs), with deletion of *Tsn1* resulting in SnToxA insensitivity (Faris et al., 2010). Similarly, the purification and subsequent isolation of the *P. nodorum* effector, SnTox1, allowed identification of the corresponding wheat sensitivity locus, *Snn1*, on the short arm of chromosome 1B (Liu et al., 2004, 2012; Cockram et al., 2015). Map-based cloning found *Snn1* to encode a wall-associated kinase (WAK), with yeast two-hybrid analysis showing the Snn1 and SnTox1 proteins interact directly *in vitro* (Shi et al., 2016b), unlike *Tsn1* and SnToxA (Faris et al., 2010). Given the nature of their corresponding wheat sensitivity loci, it is hypothesized that SnToxA and SnTox1 activate the wheat pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) and effector-triggered immunity (ETI) pathways, which for biotrophic pathogens protect against pathogen infection. However, as *P. nodorum* is a necrotrophic pathogen, the triggering of these pathways, which induce necrosis and cell death, promotes *P. nodorum* growth and propagation (Shi et al., 2016b). Wheat varieties carrying both *Tsn1* and *Snn1* show higher levels of necrosis than those varieties carrying either *Tsn1* or *Snn1* alone (Chu et al., 2010), indicating that the hijacking of both the PTI and ETI pathways for necrotrophic effector triggered susceptibility supports pathogen survival and reproduction (Shi et al., 2016b).

Characterization of a third *P. nodorum* effector, SnTox3, led to the identification of its corresponding wheat sensitivity locus, *Snn3* (more recently termed *Snn3-B1*), located on the short arm of chromosome 5B. Culture filtrate containing SnTox3 was produced using a wildtype pathogen isolate, SN15, and host sensitivity was mapped using the BR34 × Grandin wheat population, accounting for 17% of the phenotypic variation (Friesen et al., 2008). This agrees with data from the doubled haploid mapping population, Calingiri × Wyalkatchem, which identified the SnTox3 sensitivity locus *QSnb.fcu-5BS*, as well as a minor SnTox3 sensitivity QTL on the long arm of chromosome 4B, *Qsnb.cur-4BL* (Phan et al., 2016).

Understanding the effector sensitivities of wheat varieties, and the genetic determinants controlling wheat sensitivity, allows informed manipulation of alleles and germplasm within wheat breeding programs. Here, we survey a panel of 480 predominantly British winter wheat varieties for sensitivity to SnToxA, SnTox1, and SnTox3, and use this AM panel in concert with a MAGIC population to fine-map *Snn3-B1*, and to identify additional minor QTL for SnTox3 sensitivity.

Abbreviations: AM, association mapping; CIM, composite interval mapping; GWAS, genome-wide association scans; MAGIC, multi-parent advanced generation inter-cross; QTL, quantitative trait locus/loci; SMA, single marker analysis; SNB, septoria nodorum blotch; SNP, single nucleotide polymorphism.

MATERIALS AND METHODS

Wheat Germplasm and High-Density Genotyping

Two bread wheat (*T. aestivum* L.) populations were used for effector sensitivity screening and genetic mapping. The first was an AM panel, representing a diverse collection of 480 elite, predominantly British, wheat varieties drawn from historic collections and National Lists, encompassing varieties released between 1916 and 2007 (Supplementary Table S1). Of these, 420 were released or marketed within the United Kingdom; however, many of these were bred for initial release outside of the United Kingdom (data not available). The remaining 60 varieties do not have a United Kingdom Application For Protection (AFP) number, and so where country of origin information was not available, were assumed to either be non-UK, or represent accessions that predate the application process (represented in Supplementary Table S1 as accessions beginning with the prefix "U"). The majority of the AM panel represent British varieties (330 lines, 68% of the total collection), followed by 51 French (10%), 37 German (8%), and 19 Dutch varieties. The remaining 17 varieties with country information come from Belgium, Canada, Denmark, Sweden, Switzerland, and United States. For 26 varieties, country of origin was not known. The population was previously genotyped using the Illumina iSelect 90,000 feature wheat SNP array (Wang et al., 2014), resulting in 26,018 polymorphic SNPs with a minor allele frequency $\geq$ 3% (available via http://www.niab.com/pages/id/326/Resources). The second was an eight-founder MAGIC population, termed the "NIAB Elite MAGIC" population (Mackay et al., 2014), the founders of which (cvs. Alchemy, Brompton, Claire, Hereward, Rialto, Robigus, Soissons, and Xi19) were selected for their high seed yield, disease resistance, and their range of end-use qualities. The founders were intercrossed in a simple replicated funnel crossing scheme over three generations, with individuals from the eight-way families subsequently selfed over four generations through single seed descent to produce >1,000 recombinant inbred lines. A subset of these F_5 recombinant inbred lines were genotyped using the 90,000 feature SNP array detailed above, resulting in 20,643 polymorphic markers (Gardner et al., 2016). These data allowed the development of a high-resolution genetic map consisting of 18,601 markers mapped using 643 MAGIC lines (Gardner et al., 2016). The remaining 2,042 SNPs were not mappable, due largely to segregation distortion and/or dominance (Gardner et al., 2016).

Effector Protein Production, Wheat Phenotyping, and Pedigree Analysis

SnTox1 and SnTox3 were expressed in *Pichia pastoris*, as previously described (Tan et al., 2014). For SnToxA, heterologous expression was conducted in *Escherichia coli* BL21E using the pET21a expression vector as described in Tan et al. (2012). Protein preparations were desalted in 20 mM pH 7.0 sodium phosphate, freeze-dried for storage, and subsequently re-suspended in ultra-pure water and stored at 4°C prior to use. The AM and MAGIC lines were grown in 96-well trays with fine/medium compost (M3) in a heated and lit glasshouse at 20°C/17°C day/night with a 16-h photoperiod. Each line was represented by three to four replicates, and each MAGIC founder by eight replicates, with experimental design carried out using MATLAB (MATLAB, The MathWorks Inc., Natick, Massachusetts, United States) or R/blocksdesign[1]. For the AM panel, randomization was performed using a custom software routine written in the MATLAB programming environment, to include three biological replicates of each line. For the MAGIC population, the experimental design was split into four blocks, each block containing one replicate of each line and two replicates of each of the parents, with line positions randomized within each block. Therefore, a total of four biological replicates of each line and eight biological replicates of each parent were included. Infiltration on the AM and MAGIC populations was undertaken on seedlings at Zadoks growth stage (GS) 12 (Zadoks et al., 1974), as previously described (Tan et al., 2012). Briefly, a 1-ml plastic syringe was used to infiltrate approximately 50 µl of either SnToxA, SnTox1, or SnTox3 suspension into the first leaf, with the extent of the leaf infiltration region marked with a non-toxic pen. Seven days following infiltration, the plants were visually evaluated for SnTox3 effector sensitivity on a scale of 0 (insensitivity, no symptoms) to 4 (extensive necrosis; Tan et al., 2012). A water control was also used to establish a symptom baseline to evaluate possible damage due to the infiltration process. Mean sensitivity scores were calculated per line for subsequent analysis. Wheat pedigree information was obtained from public sources, and displayed using the software Helium v.1.17.08.14 (Shaw et al., 2014).

Statistical Analyses and Bioinformatics

Effector phenotypic data were analyzed using R (R Core Team, 2013) to determine mean sensitivity scores for each variety and variance within variety. For each toxin, the heritability of line means (broad sense) was calculated by first estimating components of variation from ANOVA (in the AM panel) or REML (in MAGIC) while taking into account all features of the experimental designs. Heritability was then estimated as $h^2 = \sigma^2_G/(\sigma^2_G + \sigma^2 e)$ where σ^2_G is the genetic variation between line means and $\sigma^2 e$ is the error variance appropriate to those means. Calculations were carried out in GenStat (VSN International, 2011) and the package lme4 (Bates et al., 2015) in R. GWAS using the AM panel was undertaken using the efficient mixed model association (EMMA) algorithm (Kang et al., 2008) using a compressed mixed linear model (CMLM; Zhang et al., 2010) which includes both fixed and random effects, implemented with the Genome Association and Prediction Integrated Tool (GAPIT) package (Lipka et al., 2012) in R. The genotyping and quality control of the 26,016 SNPs, and the generation of the kinship matrix, is previously described, with the majority of the population structure evident due to spring or winter seasonal growth habit (Gardner et al., submitted). For GWAS, Bonferroni corrected $P = 0.05$ and $P = 0.01$ significance thresholds were used, termed here "significant" and "highly

[1] cran.r-project.org

significant," respectively. For MAGIC, correction for multiple testing was carried out using R/qvalue using corrected threshold of 0.05. MAGIC genetic analyses were undertaken using two approaches.

(1) SMA: a simple linear model test in R/lme4 using all 20,643 SNPs. After finding of a major QTL, the analysis was repeated with the major QTL as a covariate.
(2) Haplotype analysis using a subset of 7,369 uniquely mapped SNPs from the MAGIC genetic map (Gardner et al., 2016).

Founder haplotype probabilities were computed with the "mpprob" function in R/mpMap (Huang and George, 2011), implemented in R/qtl (Broman et al., 2003), using a threshold of 0.5. QTL analysis with haplotypes was carried out (a) by linear mixed model using all mapped markers and (b) by CIM using the mpIM function in R/mpMap, with either 0, 5, or 10 covariates. The function sim.sigthr in R/mpMap was used to conduct 100 simulations using the dataset to obtain an empirical QTL significance threshold at a P threshold of 0.05. A full QTL model was then fitted with all QTL using R/fit.mpQTL. For the AM panel, the difference between homozygous marker classes was estimated as twice the linear regression coefficient from a regression of the trait on the marker classes (coded 0, 1, and 2 with 1 being the heterozygous genotype). The coefficient of determination (or r-squared) was used as a measure of the proportion of variance explained by the marker. Genetic markers were anchored using BLASTn (Altschul et al., 1990) against the wheat cv. Chinese Spring 42 IWGSC RefSeq v1.0 physical map (pre-publication data made available under the IWGSC General Data Access Agreement via https://wheat-urgi.versailles.inra.fr/), and where explicitly stated in the text, against the TGACv1 cv. Chinese Spring 42 physical map (Clavijo et al., 2017). In the case of hits of equal match on multiple homoeologs, chromosome allocation followed that assigned by the genetic map (Gardner et al., 2016), where possible. Nomenclature for the QTL discovered in this study follows that recommended by the Catalog of Gene Symbols for Wheat (McIntosh et al., 2008). Protein domains were identified using Pfam 31.0 (Finn et al., 2016).

KASP Marker Development

Single nucleotide polymorphisms were converted to the Kompetitive Allele-Specific PCR (KASP) genotyping platform (LGC Genomics, United Kingdom). SNP flanking DNA sequences were used to design KASP primers using the software PolyMarker (Ramirez-Gonzalez et al., 2015). Genomic DNA was extracted from seedling leaves harvested from a subset of the AM panel using a modified Tanksley protocol (Fulton et al., 1995), DNA concentrations determined using a Nanodrop 200 spectrophotometer (Thermo Scientific), and diluted to a final concentration of 10 ng/µl using sterile PCR-grade water. KASP genotyping was undertaken by a service provider following the manufacturer's guidelines (LGC Genomics), returned as .csv files, analyzed with SNP Viewer v.1.99[2], and compared against the corresponding SNP calls from the Illumina 90k SNP array.

[2]http://lgcgenomics.com/

RESULTS

P. nodorum Effector Sensitivity Phenotyping

The AM panel, consisting of 480 varieties and breeding lines released between 1916 and 2007, was phenotyped for sensitivity to SnToxA, SnTox1, and SnTox3 via leaf infiltration, and the severity of host response scored using a 0 (insensitive) to 4 (extensive necrosis) scale (**Figure 1** and Supplementary Table S1). Broad sense heritability for effector sensitivity was found to be highest for SnTox3 (h^2 = 0.92), followed by SnToxA (h^2 = 0.83) and SnTox1 (h^2 = 0.77). For SnToxA, a separation into insensitive/weakly sensitive (0 ≤ score < 1) and strongly sensitive (score ≥ 3) was observed, accounting for 73% (338/460 varieties) and 10% (47/460 varieties) of all lines successfully screened, respectively (**Figure 1A**). While spring cultivars represent just 6% of the complete AM panel, 23% (11/47 cultivars) of all SnToxA sensitive lines are spring cultivars, as is the oldest SnToxA sensitive variety, Garnet (released in 1926). Analysis of the wheat pedigree (**Figure 2**) shows the SnToxA sensitive varieties Axona (released in 1984) and Tonic (released in 1983) to be the most prominent in the transmission of strong sensitivity (due to the prominence of their offspring Cadenza in the wheat pedigree), accounting for 23% of varieties with a sensitivity score ≥ 3. Indeed, the four most recent SnToxA sensitive varieties

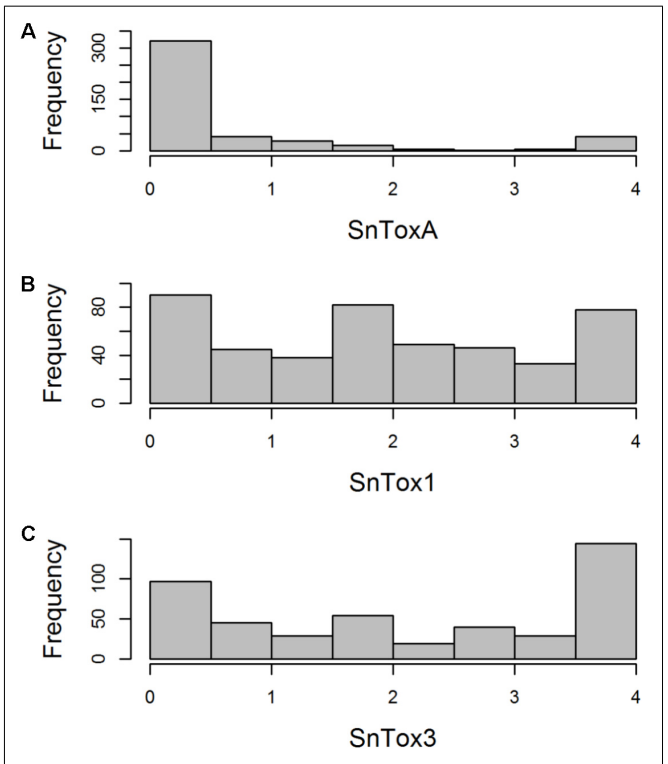


FIGURE 1 | Histogram of host sensitivity to *P. nodorum* effectors **(A)** SnToxA, **(B)** SnTox1, and **(C)** SnTox3 in the AM panel of 480 northwest European wheat varieties. Sensitivity score (0 = insensitive, 4 = sensitive) for each accession represents a mean of four biological replicates.

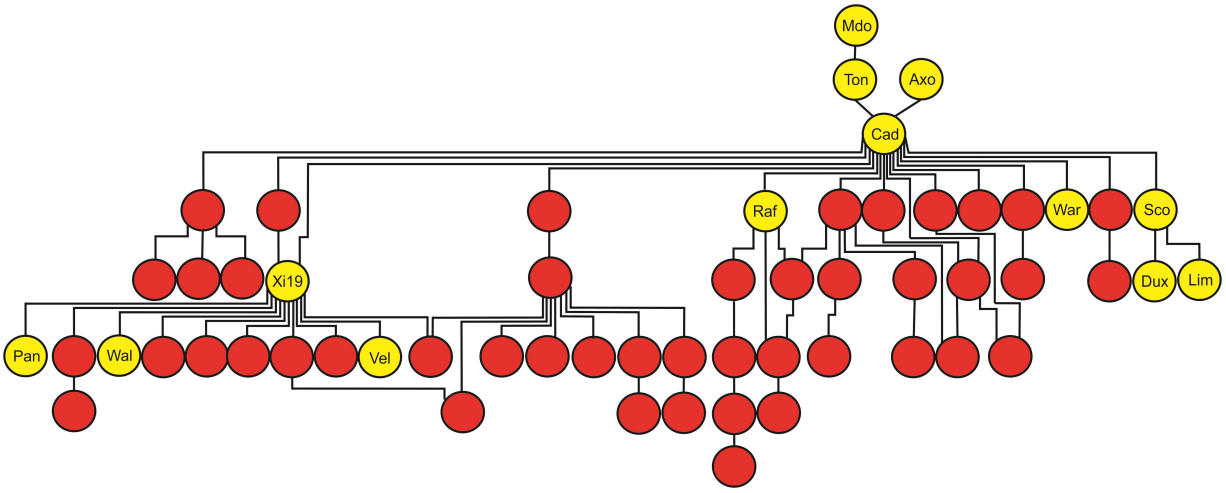


FIGURE 2 | Tracking the transmission of SnToxA sensitivity in the wheat pedigree around the Axona × Tonic pedigree. SnToxA sensitive (score ≥ 3) = yellow. SnToxA insensitive (score < 1) = red. Variety names listed only for SnToxA sensitive varieties in the pedigree: Mdo = Maris Dove, Ton = Tonic, Axo = Axona, Cad = Cadenza, Raf = Raffles, War = Warlock_24, Sco = Scorpion_25, Dux = Duxford, Lim = Limerick, Pan = Panorama, Wal = Walpole, Vel = Velocity, and Cur = KWS Curlew.

(Duxford, KWS Curlew, Limerick, and Velocity, released in 2006) all have Cadenza in their pedigree. Response to SnTox1 is more evenly spread across sensitivity classes, with 24% insensitive (111/461 varieties), 28% sensitive (130/461), and 48% (220/461) showing intermediate sensitivity (1 ≥ score < 3; **Figure 1B**). No pedigree relationships of note were observed for SnTox1 and SnTox3 sensitivity. The most common effector sensitivity identified in the AM population was to SnTox3, with 42% of varieties found to be strongly sensitive (192/457), including the recent varieties Cocoon, KWS Santiago, Orator, Rainbow, and Tuxedo, all released in 2007. Insensitivity to SnTox3 was found in 25% (114/457) of lines, including six released in 2007, while intermediate sensitivity was observed in 33% of varieties. All possible combinations of sensitivity between the three effectors were identified, with 18 varieties displaying high sensitivity (score ≥ 3) to all three effectors.

Genetic Analysis of SnTox3 Sensitivity Using the AM Panel

Initially, the precision of the AM panel was assessed empirically by undertaking GWAS for SnToxA sensitivity, known to be due to allelic variation at the gene underlying *Tsn1*. Using a data matrix of 26,018 SNPs across 480 varieties, and a Bonferroni corrected $P = 0.01$ significance threshold ($-\log_{10}P = 6.41$), GWAS identified 30 highly significant markers associated with SnToxA sensitivity (Supplementary Table S2A). These accounted for between 25% and 60% of the total variation (average 38%) with differences in score between the two homozygous classes ranging from 1.5 to 3.1 (mean 2.5). Of the eight most significant markers ($-\log_{10}P \geq 31.34$), six are located within a gene model encoding a potassium transporter (TraesCS5B01G368500), just two genes proximal to the *S/TPK-NBS-LRR* gene underlying *Tsn1* (Faris et al., 2010). Similarly, GWAS of SnTox1 sensitivity identified seven highly significant ($-\log_{10}P > 6.41$) SNPs. These accounted for between 8.3% and 14.3% of the variation (average 11.6%) with differences in score between the two homozygous classes ranging from 0.9 to 1.6 (mean 1.4). All seven SNPs were located at the *Snn1* locus, with SNP Excalibur_c21898_1423 located 25 genes distant from the *WAK* gene underlying *Snn1* in cv. Chinese Spring (TraesCS1B01G004100; Supplementary Table S2B).

Having demonstrated the utility of the AM panel, we proceeded to use SnTox3 sensitivity phenotypic data to undertake GWAS, identifying 14 significant SNPs (**Table 1A** and **Figure 3**). Of these, seven were located within a single region on the short arm of chromosome 5B in the IWGSC RefSeq v1.0 physical map. While the remaining three SNPs (wsnp_Ex_c9301_15450818, IACX7443, and Ra_c68425_1406) returned hits on unallocated chromosomes in IWGSC RefSeq v1.0, all three were localized to chromosome 5B on the TGACv1 wheat reference sequence (**Table 1A**), and have been previously allocated to the short arm of chromosome 5B by treating the SNP as a trait, and locating its estimated position by trait mapping (Gardner et al., 2016). The most significant SNPs were Excalibur_c47452_183 and GENE-3324_338 ($-\log_{10}P = 20.44$) and explained 32% of the phenotypic variation, explaining a phenotypic difference between homozygous allele classes of 2.1. Anchoring previous markers identified as flanking *Snn3-B1* to the IWGSC RefSeq v1.0 wheat genome reference sequence confirmed we had identified *Snn3-B1* in the GWAS panel. Using our peak *Snn3-B1* marker Excalibur_c47452_183 as a cofactor in GWAS found no additional significant genetic loci outside of the *Snn3-B1* region (Supplementary Table S3).

Genetic Analysis of SnTox3 Sensitivity Using MAGIC

While screening the AM panel for SnTox3 sensitivity, the eight founders of the NIAB Elite MAGIC population were found

TABLE 1 | Significant SNPs ($P \leq 0.05$, Bonferroni corrected) for SnTox3 sensitivity identified in **(A)** the AM panel and **(B)** single marker analysis (SMA) on 20,643 SNPs in the MAGIC population, listing the SNPs with effect > 1.0 (see Supplementary Table S4 for details of all 114 significant SNPs).

SNP	MAF	$-\log_{10}P$	Chr, SNP position (bp)	IWGSC RefSeq v1.0 gene model	Gene annotation
A					
Excalibur_c47452_183	0.194	23.62	5B, 6654166	TraesCS5B01G005100	Ubiquitin conjugating enzyme E2
GENE-3324_338	0.194	23.62	5B, 6647920†	TraesCS5B01G005000	Nucleotide triphosphate hydrolase
BobWhite_c4838_58	0.194	23.62	5B, 6654053	TraesCS5B01G005100	Ubiquitin conjugating enzyme
BS00091518_51	0.194	23.62	5B, 6648547	TraesCS5B01G005000	Nucleotide triphosphate hydrolase
BS00091519_51	0.194	23.62	5B, 6648567	TraesCS5B01G005000	Nucleotide triphosphate hydrolase
BS00064297_51b	0.272	16.52	5B, 6974807	TraesCS5B01G005600	Transmembrane protein, putative (DUF594)
BS00064298_51b	0.272	16.52	5B, 6974825	TraesCS5B01G005600	Transmembrane protein, putative (DUF594)
Ex_c1846_1818b	0.449	14.87	5B, 64632597	TraesCS5B01G059000	Transmembrane protein, putative (DUF594)
Ex_c1846_1818a	0.271	13.46	5B, 64736555	TraesCS5B01G059000	Transmembrane protein, putative (DUF594)
RAC875_c7582_680	0.124	7.88	5B, 2058821	TraesCS5B01G002000LC	None
Ku_c10387_272	0.124	7.88	5B, 232228	TraesCS5B01G000600	Microtubule-associated protein 70-2
RAC875_c39204_91	0.157	7.11	5B, 6852650	None	None
BS00064297_51a	0.156	7.02	5B, 6974807	TraesCS5B01G005600	Transmembrane protein, putative (DUF594)
BS00064298_51a	0.156	7.02	5B, 6974825	TraesCS5B01G005600	Transmembrane protein, putative (DUF594)
B					
Excalibur_c47452_183		15.7‡	5B, 6654166	TraesCS5B01G005100	Ubiquitin conjugating enzyme E2
GENE-3324_338		15.7‡	5B, 6647920†	TraesCS5B01G005000	Nucleotide triphosphate hydrolase
BobWhite_c4838_58		15.7‡	5B, 6654053	TraesCS5B01G005100	Ubiquitin conjugating enzyme E2
BS00091518_51		15.7‡	5B, 6648547	TraesCS5B01G005000	Nucleotide triphosphate hydrolase
BS00091519_51		15.7‡	5B, 6648567	TraesCS5B01G005000	Nucleotide triphosphate hydrolase

MAF, minor allele frequency (AM panel only); Chr, chromosome; U, chromosome unknown. †*While chromosome 5A homoeolog represented a better BLASTn hit, details for the 5B homoeologs are presented here, as this SNP has previously been shown to be located on chromosome 5B by using the SNP as a trait, and localizing to a chromosome by trait mapping (Gardner et al., 2016).* ‡*P-values* $< 2.2^{-16}$*, so* $-\log_{10}P$ *set arbitrarily to > 15.7 here.*

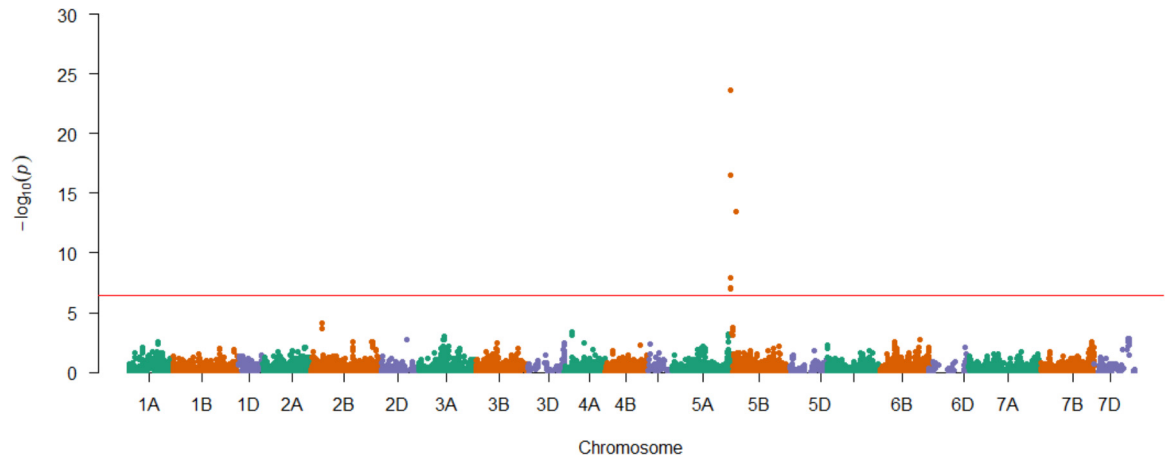


FIGURE 3 | Genetic mapping of SnTox3 sensitivity in the AM panel. SNPs are ordered according to the genetic map (Gardner et al., 2016). Significant SNPs that were unmapped in the genetic map (Gardner et al., 2016) are included on the relevant chromosome (5B) via their physical map positions (IWGSC RefSeq v1.0), relative to those of the genetically mapped chromosome 5B SNPs. The Bonferroni corrected $P = 0.01$ significance threshold is indicated in red ($-\log_{10}P = 6.41$).

to contrast for SnTox3 sensitivity (Supplementary Table S1). Accordingly, 643 lines of the MAGIC population (3–4 reps/line), as well as the eight founders (8 reps/line), were subsequently phenotyped for SnTox3 sensitivity. The sensitivity scores of the parents ranged from 0 (Alchemy, Claire, Robigus) to 4 (Hereward, Rialto, Soissons, Xi19), with Brompton displaying intermediate sensitivity (1.93). SnTox3 sensitivity scores in the MAGIC progeny ranged from 0 to 4, with the majority of lines displaying high sensitivity (score $\geq$ 3.5, 42.5% lines) or no sensitivity (score $\leq$ 0.5, 27.5% lines; Supplementary Figure S1). Heritability for SnTox3 sensitivity in the MAGIC population was calculated to be $h^2 = 0.95$. Initially, the 643 MAGIC lines along with the 20,643 mapped and unmapped SNPs were used for SMA using a simple linear model test,

identifying 114 significant ($P < 0.05$) markers (Supplementary Table S4). Of these, the five most significant SNPs, with predicted allelic effects > 1, are the same as the five most significant markers identified in the AM panel (Excalibur_c47452_183, GENE-3324_338, BobWhite_c4838_58, BS00091518_51, and BS00091519_51; **Table 1B**). SMA identified four additional QTL. The first was on chromosomes 2B (termed here *QTox3.niab-2B.1*, $P = 0.023$), located at 356.66 cM by SNP Kukri_c9898_1766. The second, at 40.61 cM on chromosome 4D (*QTox3.niab-4B.1*, $P = 0.037$), was identified by SNP BS00036421_51, the third (*QTox3.niab-6B.1*) was identified by three SNPs on chromosome 6B, with wsnp_Ku_c2119_4098330 showing the highest significance ($P = 0.003$), and the fourth (*QTox3.niab-7B.1*) was identified by four chromosome 7B SNPs, with BS00022127_51 showing the highest significance ($P = 0.038$). Using the most significant peak SNP GENE_3324_338 as a covariate in SMA analysis did not identify any additional genetic loci, and resulted in the disappearance of all four minor QTL.

Additionally, CIM using 0, 5, and 10 covariates was undertaken, using the 7,369 uniquely mapped SNPs from the MAGIC genetic map (Gardner et al., 2016). As well as identifying *Snn3-B1* on chromosome 5B (*P*-values for all analyses $\leq 2.2^{-16}$, accounting for ≥16.95% of the phenotypic variation), we again detected *QTox3.niab-2B.1* ($P = 0.047$, 2.3% variation explained, detected with 0 covariates only) and *QTox3.niab-7B.1* ($P = 0.025$, 1.6% variation explained, SNP detected with 0 covariates only). In addition, three further QTL were discovered, distinct to those identified by SMA (**Table 2**). The first, *QTox3.niab-2A.1*, mapped to chromosome 2A at 234.62 cM (SNPs BS00070979_51 and Excalibur_c20478_641, positioned at ∼758 Mb) with a *P*-value = 9.17^{-7}, and explained 6.0% of the variation with 0 covariates. The second, *QTox3.niab-3B.1*, was located on 3B at 84.11 cM ($P = 48.51^{-4}$, SNPs wsnp_Ex_c11246_18191331, wsnp_Ex_c22401_31592784, ∼68 Mb), and explained 3.1% of the variation (only found with 5 or 10 covariates). Finally, *QTox3.niab-6A.1* at 65.6 cM on chromosome 6A explained 4.2% of the variation with 0 covariates ($P = 8.51^{-4}$, SNPs BobWhite_c13839_135 and IACX7801, ∼22 Mb; Supplementary Table S7).

Development of KASP Genetic Markers for *Snn3-B1*

The peak SNP identified in both the AM panel by GWAS and MAGIC population by SMA was Excalibur_c47452_183. This marker was selected for conversion from the 90k SNP array to the KASP genotyping platform, a single-plex technology that allows flexible, low-cost use for marker-assisted breeding and research. Primers were designed and tested on a subset of 95 varieties from the AM panel (**Figure 4** and Supplementary Table S5). Comparison of Excalibur_c47452_183 allele calls from KASP genotyping with those returned by the 90k array genotyping of the AM panel found perfect correspondence between the two, indicating robust conversion to the KASP platform (Supplementary Table S5). This SNP provides good, but not perfect, prediction of SnTox3 sensitivity phenotype in the AM panel (**Table 3** and Supplementary Table S1).

TABLE 2 | Significant SnTox3 sensitivity QTL identified by CIM in the MAGIC population, incorporating 0, 5, and 10 covariates in the analysis (labeled below as cov 0, cov 5, and cov 10, respectively).

Cov 0			Cov 5			Cov 10		
Chr, cM	***P*-value, % Va**	**SNPs flanking QTL peak**	**Chr, cM**	***P*-value, % Va**	**SNPs flanking QTL peak**	**Chr, cM**	***P*-value, % Va**	**SNPs flanking QTL peak**
2A, 234.62	9.17^{-7}, 6.0	BS00070979_51, Excalibur_c20478_641	2A, 235.12	1.18^{-5}, 4.9	Excalibur_c20478_641, Tdurum_contig56321_232	2A, 234.62	3.22^{-6}, 2.8	BS00070979_51, Excalibur_c20478_641
2B, 379.61	4.7^{-3}, 2.3	BS00064483_51, Kukri_c1526_666						
			3B, 84.11	8.51^{-4}, 3.1	wsnp_Ex_c11246_18191331, wsnp_Ex_c22401_31592784	3B, 84.11	7.96^{-4}, 3.1	wsnp_Ex_c11246_18191331, wsnp_Ex_c22401_31592784
5B, 1.27	3.33^{-16}, 15.5	BS00015136_51, GENE-3277_145	5B, 3.29	0, 17.0	BS00025784_51, BS00065732_51	5B, 3.29	0*, 17.0	BS00025784_51, BS00065732_51
6A, 65.6	6.16^{-5}, 4.2	BobWhite_c13839_135, IACX7801				6A, 65.6	3.53^{-4}, 3.2	BobWhite_c13839_135, IACX7801
7B, 9.43	0.025, 1.6	BS00022127_51, Kukri_c67849_109						

*% Va, percent variation explained. QTL identified by two or more covariate analysis methods are reported in the manuscript, using P-values and SNP information from the analysis with the least number of covariates. Chromosome (Chr) and cM positions are from the NIAB Elite MAGIC genetic map (Gardner et al., 2016). * <2.2–16.*

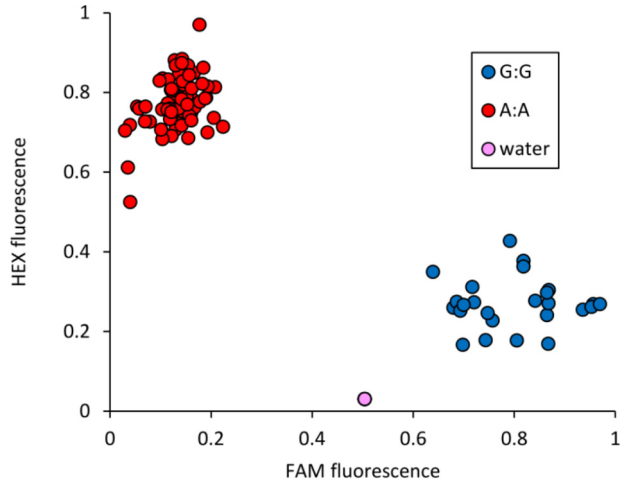


FIGURE 4 | Validation of KASP marker Excalibur_c47452_183 (SNP assayed = A/G), closely linked to *Snn3-B1*. Allele-specific primer A: 5′-GAAGGTCGGAGTCAACGGATTaaggctggctggcgagtA-3′. Allele-specific primer B: 5′-GAAGGTGACCAAGTTCATGCTaaggctggctggcgagtG-3′. Common primer: 5′-tgaggggcatcgaatcG-3′. Tails for ligation of fluorophores are included in uppercase at the 5′ end of the allele-specific primers.

TABLE 3 | Allele call for SNP marker Excalibur_c47452_183 versus SnTox3 sensitivity score in the AM panel.

	SnTox3 sensitivity < 4	SnTox3 sensitivity = 4	Total no. of varieties†
Allele A:A	317	43	360
Allele G:G	6	91	97
Total no. of varieties†	323	134	457

†*Accessions for which phenotypic data were available.*

Analysis of the *Snn3-B1* Physical Region

To investigate gene content at the *Snn3-B1* locus on chromosome 5B, the sequences containing the most significant SNPs identified in the AM panel ($-\log_{10}P > 16$, seven SNPs) and the MAGIC population by SAM ($-\log_{10}P > 16$, effect > 1, five SNPs) were aligned to the IWGSC RefSeq v1.0 wheat reference genome. The seven SNPs from the AM panel delineated a physical region of 326 kb (∼6.648–6.975 Mb), while the five MAGIC SNPs delineated physical regions of 6.2 kb (6.648–6.654 Mb), located within the physical interval as defined in the AM panel. The 326-kb region was predicted to contain seven gene models, representing three high-confidence and four low-confidence gene models (Supplementary Table S6). Two markers lie within a gene model encoding a ubiquitin-conjugating enzyme: Excalibur_c47452_183 is in the 5′ untranslated region, while BobWhite_c4838_58 is a synonymous SNP located within exon 6. Three SNPs lie within gene model TraesCS5B01G005000, a P-loop containing nucleoside triphosphate hydrolases superfamily protein: GENE-3324_338 is located in intron 2, BS00091518_51 results in a synonymous substitution in exon 3, and BS00091519_51 is predicted to result in a G→D substitution at amino acid residue 675 (G675/D), outside of the 50s ribosome-binding GTPase domain. Finally, two SNPs are located in gene model TraesCS5B01G005600, encoding a putative transmembrane protein: BS00064297_51b and BS00064298_51b both represent non-synonymous mutations (L474/P and Q480/R, respectively), neither of which are predicted to lie within a known protein domain.

DISCUSSION

Effector Sensitivity in European Wheat Germplasm and Its Relevance to SNB

Septoria nodorum blotch is a major disease of wheat in many growing areas, with field resistance based on multiple minor effect genes. Identification of necrotrophic effectors in *P. nodorum* provided resources with which to dissect host resistance into its constituent parts (Cockram et al., 2015) and study their interactions (Phan et al., 2016). Here we used effector screening to determine sensitivities in 480 predominantly British winter wheat varieties. The frequency of SnTox1 sensitive varieties was 28%, broadly comparable to that found for Scandinavian varieties (12%, Ruud et al., 2017a) and global collections (16%, Shi et al., 2016a), but contrasts notably against a recent screen of Australian varieties (72%, Tan et al., 2014). SnTox3 sensitivity frequency here (42%) was similar to that reported in Scandinavian germplasm (55%, Ruud et al., 2017a). Sensitivity to SnToxA was 10% in the predominantly British winter wheat germplasm collection screened here, which is notably lower than that reported in other wheat germplasm collections, e.g., 45% in Scandinavian varieties (Ruud et al., 2017a) and 65% in Western Australian wheat (Waters et al., 2011). SnToxA sensitivity was found to be present at a relatively high frequency in the spring wheat varieties in our panel. The spring and winter wheat breeding pools are relatively separate, possibly explaining the observed frequency differences in SnToxA sensitive and insensitive alleles between the two groups. Indeed, where sensitive alleles appear in the winter genepool, this can often be tracked via the parentage of the genetically spring cultivar Cadenza, a prominent variety in the British pedigree. *Tsn1* is located on the long arm of chromosome 5B, <30 Mb from the vernalization gene *VRN-B1*, a major gene influencing winter or spring growth habit (Cockram et al., 2007). Therefore, it is likely that the partitioning of *Tsn1* alleles is influenced by linkage to winter or spring alleles at *VRN-B1*.

Genetic Mapping of SnTox3 Sensitivity

Previous studies have mapped *Snn3-B1* to the short arm of chromosome 5B (Friesen et al., 2008; Phan et al., 2016; Shi et al., 2016a; Ruud et al., 2017b). Here, genetic analysis in the AM and MAGIC populations identified *Snn3-B1* as representing the most significant genetic determinant of SnTox3 sensitivity in British germplasm, with the highly significant markers ($-\log_{10}P > 10$) delimiting a 327-kb interval containing seven genes. This was achieved without the need to develop populations specifically to investigate SnTox3 sensitivity, demonstrating the efficacy of using genetic resources such as AM panels and MAGIC populations for rapid genetic dissection of target traits (Cockram and Mackay,

2018). The relatively small genetic interval determined here will allow reverse genetic approaches such as genome editing and Targeting Induced Local Lesions IN Genomes (TILLING, McCallum et al., 2000) to be undertaken to help identify the gene underlying *Snn3-B1*. However, it should be noted that given we found the cultivar Chinese Spring from which the wheat reference genome sequence is derived to be insensitive to SnTox3, it is possible that the gene underlying *Snn3-B1* is deleted or degraded to such an extent that it is not predicted as a gene model. Indeed, this was the case for the gene underlying the SnToxA sensitivity locus *Tsn1*, which was found to be absent in insensitive varieties, including Chinese Spring (Faris et al., 2010).

Minor QTL for SnTox3 sensitivity have only previously been reported on chromosome 4B (Phan et al., 2016). The six QTL identified here in the MAGIC population, on chromosomes 2A, 2B, 3B, 4D, 6A, and 7B, are therefore novel. When these are compared with known QTL for SNB field resistance and *P. nodorum* juvenile resistance studies, possible overlap can be identified for the 2A MAGIC QTL *QTox3.niab-2A.1*. This likely corresponds to *Qsnb.cur-2AS.1* (Supplementary Table S7), controlling seedling SNB sensitivity using *P. nodorum* isolate SN15, as well as knock-out strains of this isolate lacking SnTox1 (*tox1–6*), a triple knock-out strain lacking SnToxA, SnTox1, and SnTox3 (*toxa13*), and seedling inoculation with culture filtrate from isolate *toxa13* (Phan et al., 2016), and to the SNB QTL identified in the Arina × Forno population via common marker gwm372 (Abeysekara et al., 2009). Similarly, the SnTox3 sensitivity QTL identified here on chromosome 3B (*QTox3.niab-3B.1*) appears to correspond to a QTL for adult plant SNB resistance in the wheat SHA3/CBRD x Naxos grown under field conditions in Norway (Ruud et al., 2017b). The creation of near isogenic lines for such QTL would allow further characterization of their effects, and potentially, isolation of their underlying genes. The MAGIC population consists of inbred lines genotyped at the F_5 generation, with each line expected to contain ∼2% heterozygosity, allowing development of heterogeneous inbred families (HIFs) to rapidly create near isogenic lines through selfing (Tuinstra et al., 1997). As MAGIC F_5 lines heterozygous across each of the minor QTL are available, it should now be possible to rapidly create precise genetic materials with which to investigate their effects in isolation. We note that while MAGIC QTL analysis using CIM allowed minor QTL to be detected, in comparison to MAGIC SMA analyses (and GWAS analysis in the AM panel), it did not accurately locate *Snn3-B1*. This is due to a 5BS/7BS translocation (Badaeva et al., 2007) that is known to segregate in the MAGIC population, with the resulting segregation distortion preventing genetic mapping of markers close to the translocation breakpoint (Gardner et al., 2016). As *Snn3-B1* is close to this breakpoint, the absence of the most closely linked SNPs in the genetic map prevents accurate mapping via CIM. In contrast, SMA analysis in the MAGIC population does not require markers to be genetically mapped, highlighting the importance of using both analysis methods when undertaking QTL analysis. Indeed, in addition to *Snn3-B1*, SMA and CIM both identified additional QTL, two of which were shared and two or three of which were private to each analysis method.

Analysis of the *Snn3-B1* Physical Region

The physical region, as defined by the most significant SNPs identified in the AM panel and MAGIC population, was predicted to contain seven gene models. Gene model TraesCS5B01G005000 (containing SNPs GENE-3324_338, BS00091518_51, and BS00091519_51) is similar to *YELLOW LEAF 1/BRASSINAZOLE INSENSITIVE PALE GREEN 2* (*BPG2*), involved in the accumulation of chloroplast proteins and the salt stress response pathway in Arabidopsis (Li et al., 2016; Qi et al., 2016). The SNP that had the largest phenotypic effect from the MAGIC SMA analysis was located within gene model TraesCS5B01G005100, which encoded a "ubiquitin-conjugating enzyme E2". This class of genes has been shown to regulate plant disease resistance, both positively and negatively. Examples include the U-box type E3 ubiquitin ligase, CMPG1, that regulates immunity in multiple plant species (González-Lamothe et al., 2006), SPL11, a negative regulator of cell death in rice (Zeng et al., 2004), and Plant U-box 22 (PUB22), PUB23, and PUB24, that negatively regulate PTI in Arabidopsis (Trujillo et al., 2008). Genes TraesCS5B01G005200, TraesCS5B01G005300, and TraesCS5B01G005400 all showed sequence similarity to protein kinases, a class of genes known to play a role in disease resistance (Xia, 2004). However, protein kinase domains were only predicted within the amino acid sequence of TraesCS5B01G005400. Finally, gene models TraesCS5B01G005500 and TraesCS5B01G005600 (containing SNPs BS00064297_51b and BS00064298_51b) both encode predicted transmembrane proteins, with BLASTn matches ($\leq 7e^{-66}$) to single, unannotated genes in rice and brachypodium. Further work is needed to investigate whether any of these genes underlie *Snn3-B1*.

The Use of Effector Sensitivity Loci for Wheat Research and Breeding

SnTox3 sensitivity and disease susceptibility had previously been reported to be poorly correlated (Friesen et al., 2009), only accounting for a significant portion of disease phenotype in adult plants segregating for sensitivity alleles at *Snn3-B1*, *Tsn1*, and *Snn2* – although when infected with *P. nodorum* isolates lacking *SnToxA*. This is consistent with the notion that the SnToxA–*Tsn1* interaction is epistatic to the SnTox3–*Snn3-B1* interaction (Friesen et al., 2008). However, more recently, the *Snn3-B1* locus has been identified in QTL analysis of adult plant field resistance to SNB in northwestern Europe (Waters et al., 2011; Lu and Lillemo, 2014; Ruud et al., 2017b). It is thought that *SnTox1* expression inhibits the transcription of *SnTox3* (Phan et al., 2016). This may explain the reason that while gene-for-gene interactions are readily identified via effector infiltration, their interactions are not always additive. However, recent work has found infiltration of wheat seedlings with culture filtrate using SnTox3 positive *P. nodorum* isolates resulted in a necrotic phenotype

on wheat containing *Snn3-B1*, irrespective of the presence of *SnTox1* in the pathogen (Ruud et al., 2017b). Nevertheless, the differing associations between disease susceptibility and effector sensitivity will likely depend on the effectors present in regional pathogen populations, the interactions between these effectors, and the alleles present at their corresponding host sensitivity loci. To allow rapid selection for allelic state at *Snn3-B1*, we develop a co-dominant KASP genetic marker closely linked to the locus for use within wheat breeding programs. The marker represents a useful tool for marker-assisted selection for SNB, given the proven association between *Snn3-B1* and SNB resistance and that the marker is able to robustly call alleles. Despite the relatively simple Mendelian control of the trait in bread wheat, however, this marker is not a perfect predictor of SnTox3 sensitivity in the 457 phenotyped accessions in the AM panel – most notably, the 43 highly SnTox3 sensitive varieties that carry A:A alleles at SNP Excalibur_c47452_183 (**Table 3** and Supplementary Table S1). This is similar to the results of other studies that have attempted to identify diagnostic markers for SnTox3 sensitivity (Shi et al., 2016a; Phan et al., 2018). This observation could be due to a number of reasons, including one or a combination of the following: insufficient marker saturation, multiple alleles at the *Snn3-B1* locus, control by copy number variation, or the effect of minor QTL. Indeed, SnTox3 sensitivity in the AM panel shows more of a quantitative distribution, in contrast to the qualitative phenotypic distribution found for SnToxA sensitivity (**Figure 1**). Nevertheless, the KASP marker for SNP Excalibur_c47452_183 developed here will be of use in tracking SnTox3 sensitivity alleles where the sensitivity of the founders is known (e.g., Supplementary Table S1), and to help further narrow the *Snn3-B1* genetic interval. For example, given that nearly all varieties phenotyped that carry G:G alleles, with few exceptions, are highly sensitive (score $>$ 3) for SnTox3, this marker could be used to remove the majority of highly sensitive varieties from a breeding program.

AUTHOR CONTRIBUTIONS

RD, LB, GM, and GR undertook research. RD, KG, NG, CCM, IM, and JC analyzed the data. EF, JC, KG, HP, IM, and RO provided scientific input and project resources. JC, K-CT, NG, and RO provided supervision and project management. RD and JC wrote the manuscript. All authors reviewed the manuscript.

FUNDING

This work was funded by Biotechnology and Biological Sciences Council (BBSSRC) grant BB/N00518X/1, BBSRC Ph.D. grant to RD, and by Curtin University – Grains Research and Development Corporation bilateral research grant CUR00023. Joint coordination and planning of project activities by JC and RO was aided by networking activities funded under the COST Action "SUSTAIN".

ACKNOWLEDGMENTS

The authors would like to thank the International Wheat Genome Sequencing Consortium for prepublication access to IWGSC RefSeq v1.0.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.00881/full#supplementary-material

REFERENCES

Abeysekara, N. S., Friesen, T. L., Keller, B., and Faris, J. D. (2009). Identification and characterization of a novel host-toxin interaction in the wheat-*Stagonospora nodorum* pathosystem. *Theor. Appl. Genet.* 120, 117–126. doi: 10.1007/s00122-009-1163-6

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Badaeva, E. D., Dedkova, O. S., Gay, G., Pukhalskyi, V. A., Zelenin, A. V., Bernard, S., et al. (2007). Chromosomal rearrangements in wheat: their types and distribution. *Genome* 50, 907–926. doi: 10.1139/G07-072

Ballance, G. M., Lamari, L., and Bernier, C. C. (1989). Purification and characterisation of host-selective necrosis toxin from *Pyrenophora tritici-repentis*. *Physiol. Mol. Plant Pathol.* 35, 203–213. doi: 10.1016/0885-5765(89)90051-9

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bhathal, J. S., and Loughman, R. (2003). Yield reduction in wheat in relation to leaf disease from yellow (tan) spot and septoria nodorum blotch. *Eur. J. Plant Pathol.* 109, 435–443. doi: 10.1023/A:1024277420773

Broman, K. W., Wu, H., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19, 889–890. doi: 10.1093/bioinformatics/btg112

Chu, C.-G., Faris, J. D., Xu, S. S., and Friesen, T. L. (2010). Genetic analysis of disease susceptibility contributed by the compatible *Tsn1-SnToxA* and *Snn1-SnTox1* interactions in the wheat-*Stagonospora nodorum* pathosystem. *Theor. Appl. Genet.* 120, 1451–1459. doi: 10.1007/s00122-010-1267-z

Clavijo, B. J., Venturini, L., Schudoma, C., Accinelli, G. G., Kaithakottil, G., Wright, J., et al. (2017). An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Res.* 27, 885–896. doi: 10.1101/gr.217117.116

Cockram, J., Jones, H., Leigh, F. J., O'Sullivan, D., Powell, W., Laurie, D. A., et al. (2007). Control of flowering time in temperate cereals: genes, domestication and sustainable productivity. *J. Exp. Bot.* 58, 1231–1244. doi: 10.1093/jxb/erm042

Cockram, J., and Mackay, I. J. (2018). "Genetic mapping populations for conducting high resolution trait mapping in plants," in *Advances in Biochemical Engineering Biotechnology*, ed. R. Varshney (Berlin: Springer International Publishing).

Cockram, J., Scuderi, A., Barber, T., Furuki, E., Gardner, K. A., Gosman, N., et al. (2015). Fine-mapping the wheat *Snn1* locus conferring sensitivity to the *Parastagonospora nodorum* necrotrophic effector SnTox1 using an eight founder multi-parent advanced generation intercross population. *G3* 5, 2257–2266. doi: 10.1534/g3.115.021584

Faris, J. D., Anderson, J. A., Francl, L. J., and Jordahl, J. G. (1996). Chromosomal location of gene conditioning insensitivity in wheat to a necrosis-inducing culture filtrate from *Pyrenophora tritici-repentis*. *Phytopathology* 86, 459–463. doi: 10.1094/Phyto-86-459

Faris, J. D., Zhang, Z., Lu, H., Reddy, L., Clouter, S., Fellers, J. P., et al. (2010). A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13544–13549. doi: 10.1073/pnas.1004090107

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam proteins database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344

Friesen, T. L., Ali, S., Klein, K., and Rasmussen, J. B. (2005). Population genetic analysis of a global collection of *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. *Phytopathology* 95, 1144–1150. doi: 10.1094/PHYTO-95-1144

Friesen, T. L., Chu, C.-G., Liu, Z. H., Xu, S. S., Halley, S., and Faris, J. D. (2009). Host-selective toxins produced by *Stagonospora nodorum* confer disease susceptibility in adult wheat plants under Weld conditions. *Theor. Appl. Genet.* 118, 1489–1497. doi: 10.1007/s00122-009-0997-2

Friesen, T. L., Friesen, Z., Solomon, P. S., Oliver, R. P., and Faris, J. D. (2008). Characterization of the interaction of a novel *Stagonospora nodorum* host-selective toxin with a wheat susceptibility gene. *Plant Physiol.* 146, 682–693. doi: 10.1104/pp.107.108761

Friesen, T. L., Meinhardt, S. W., and Faris, J. D. (2007). The *Stagonospora nodorum*-wheat pathosystem involves multiple proteinaceous host-selective toxins and corresponding host sensitivity genes that interact in an inverse gene-for-gene manner. *Plant J.* 51, 681–692. doi: 10.1111/j.1365-313X.2007.03166.x

Fulton, T. M., Chunwongse, T., and Tanksley, S. D. (1995). Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Mol. Biol. Rep.* 13, 207–209. doi: 10.1007/BF02670897

Gardner, K. A., Wittern, L. M., and Mackay, I. J. (2016). A highly recombined, high-density, eight founder wheat MAGIC map reveals extensive segregation distortion and genomic locations of introgression segments. *Plant Biotechnol. J.* 14, 1406–1417. doi: 10.1111/pbi.12504

González-Lamothe, R., Tsitsigiannis, D. I., Ludwig, A. A., Panicot, M., Shirasu, K., and Jones, J. D. (2006). The U-box protein CMPG1 is required for efficient activation of defense mechanisms triggered by multiple resistance genes in tobacco and tomato. *Plant Cell* 18, 1067–1083. doi: 10.1105/tpc.106.040998

Huang, B. E., and George, A. W. (2011). R/mpMap: a computational platform for the genetic analysis of multiparent recombinant inbred lines. *Bioinformatics* 27, 727–729. doi: 10.1093/bioinformatics/btq719

Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., et al. (2008). Efficient control of population structure in model organism association mapping. *Genetics* 178, 1709–1723. doi: 10.1534/genetics.107.080101

Li, P.-C., Huang, J.-G., Yu, S.-W., Li, T.-T., Sun, P., et al. (2016). *Arabidopsis* YL1/BPG2 is involved in seedling shoot response to salt stress through ABI4. *Sci. Rep.* 3:30163. doi: 10.1038/srep30163

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444

Liu, Z., Faris, J. D., Oliver, R. P., Tan, K.-C., Soloman, P. S., McDonald, M. C., et al. (2009). SnTox3 acts in effector triggered susceptibility to induce disease on wheat carrying the *Snn3* gene. *PLoS Pathog.* 5:e1000581. doi: 10.1371/journal.ppat.1000581

Liu, Z., Friesen, T. L., Ling, H., Meinhardt, S. W., Oliver, R. P., Rasmussen, J. B., et al. (2006). The *Tsn1*-ToxA interaction in the wheat-*Stagonospora nodorum* pathosystem parallels that of the wheat-tan spot system. *Genome* 49, 1265–1273. doi: 10.1139/g06-088

Liu, Z., Zhang, Z., Faris, J. D., Oliver, R. P., Syme, R., McDonald, M. C., et al. (2012). The cysteine rich necrotrophic effector SnTox1 produced by *Stagonospora nodorum* triggers susceptibility of wheat lines harbouring *Snn1*. *PLoS Pathog.* 8:e1002467. doi: 10.1371/journal.ppat.1002467

Liu, Z. H., Faris, J. D., Meinhardt, S. W., Ali, S., Rasmussen, J. B., and Friesen, T. L. (2004). Genetic and physical mapping of a gene conditioning sensitivity in wheat to a partially purified host-selective toxin produced by *Stagonospora nodorum*. *Phytopathology* 94, 1056–1060. doi: 10.1094/PHYTO.2004.94.10.1056

Lu, Q., and Lillemo, M. (2014). Molecular mapping of adult plant resistance to *Parastagonospora nodorum* leaf blotch in bread wheat lines 'Shanghai-3/Catbird' and 'Naxos'. *Theor. Appl. Genet.* 127, 2635–2644. doi: 10.1007/s00122-014-2404-x

Mackay, I. J., Bansept-Basler, P., Barber, T., Bentley, A. R., Cockram, J., Gosman, N., et al. (2014). An eight-parent multiparent advanced generation inter-cross population for winter-sown wheat: creation, properties, and validation. *G3* 4, 1603–1610. doi: 10.1534/g3.114.012963

McCallum, C. M., Comai, L., Greene, E. A., and Henikoff, S. (2000). Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol.* 123, 439–442. doi: 10.1104/pp.123.2.439

McIntosh, R. A., Yamazaki, Y., Dubcovsky, J., Rogers, J., Morris, C., Somers, D. J., et al. (2008). "Catalogue of gene symbols for wheat," in *Proceedings of the 11th International Wheat Genetics Symposium,* San Diego, CA, 24–29.

Oliver, R. P., Friesen, T. L., Faris, J. D., and Solomon, P. S. (2012). *Stagonospora nodorum*: from pathology to genomics and host resistance. *Annu. Rev. Phytopathol.* 50, 23–43. doi: 10.1146/annurev-phyto-081211-173019

Oliver, R. P., and Solomon, P. S. (2010). New developments in pathogenicity and virulence of necrotrophs. *Curr. Opin. Plant Biol.* 13, 415–419. doi: 10.1016/j.pbi.2010.05.003

Phan, H., Rybak, K., Bertazzoni, S., Furuki, E., Dinglasan, E., Hickey, L. T., et al. (2018). Novel sources of resistance to *Septoria nodorum* blotch in the Vavolov wheat collection identified by genome-wide association studies. *Theor. Appl. Genet.* 131, 1223–1238. doi: 10.1007/s00122-018-3073-y

Phan, H., Rybak, K., Furuki, E., Breen, S., Solomon, P. S., Oliver, R. P., et al. (2016). Differential effector gene expression underpins epistasis in a plant fungal disease. *Plant J.* 87, 343–354. doi: 10.1111/tpj.13203

Qi, Y., Zhao, J., An, R., Zhang, J., Liang, S., Shao, J., et al. (2016). Mutations in circularly permuted GTPase family genes *AtNOA1/RIF1/SVR10* and *BPG2* supress *var2*-mediated leaf variegation in *Arabidopsis thaliana*. *Photosynth. Res.* 127, 355–367. doi: 10.1007/s11120-015-0195-9

Quaedvlieg, W., Verkley, G. J., Shin, H. D., Barreto, R. W., Alfenas, A. C., Swart, W. J., et al. (2013). Sizing up septoria. *Stud. Mycol.* 75, 307–390. doi: 10.3114/sim0017

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R foundation for statistical computing.

Ramirez-Gonzalez, R. H., Uauy, C., and Caccamo, M. (2015). PolyMarker: a fast polyploid primer design pipeline. *Bioinformatics* 31, 2038–2039. doi: 10.1093/bioinformatics/btv069

Ruud, A. K., Dieseth, J. A., and Lillemo, M. (2017a). Effects of three *Parastagonospora nodorum* necrotic effectors on spring wheat under Norwegian field conditions. *Crop Sci.* 58, 159–168. doi: 10.2135/cropsci2017.05.0281

Ruud, A. K., Windju, S., Belova, T., Friesen, T. L., and Lillemo, M. (2017b). Mapping of SnTox3–*Snn3* as a major determinant of field susceptibility to *Septoria nodorum* leaf blotch in the SHA3/CBRD × Naxos population. *Theor. Appl. Genet.* 130, 1361–1374. doi: 10.1007/s00122-017-2893-5

Shaw, P. D., Kennedy, J., Graham, M., Milne, I., and Marshall, D. F. (2014). Helium: visualization of large scale plant pedigrees. *BMC Bioinformatics* 15:259. doi: 10.1186/1471-2105-15-259

Shi, G., Zhang, Z., Friesen, T. L., Bansal, U., Cloutier, S., Wicker, T., et al. (2016a). Marker development, saturation mapping, and high-resolution mapping of the *Septoria nodorum* blotch susceptibility gene *Snn3–B1* in wheat. *Mol. Genet. Genomics* 291, 107–119. doi: 10.1007/s00438-015-1091-x

Shi, G., Zhang, Z., Friesen, T. L., Raats, D., Fahima, T., Brueggeman, R. S., et al. (2016b). The hijacking of a receptor kinase-driven pathway by a wheat fungal pathogen leads to disease. *Sci. Adv.* 2:e1600822.

Solomon, P. S., Lowe, R. G., Tan, K.-C., Waters, O. D., and Oliver, R. P. (2006). *Stagonospora nodorum*: cause of stagonospora nodorum blotch of wheat. *Mol. Plant Pathol.* 7, 147–156. doi: 10.1111/j.1364-3703.2006.00326.x

Tan, K. C., Ferguson-Hunt, M., Rybak, K., Walters, O. D., Stanley, W. A., Bond, C. S., et al. (2012). Quantitative variation in effector activity of ToxA isoforms from *Stagonospora nodorum* and *Pyrenophora tritici-repentis*. *Mol. Plant Microbe Interact.* 25, 515–522. doi: 10.1094/MPMI-10-11-0273

Tan, K.-C., Walters, O. D. C., Rybak, K., Antoni, E., Furuki, E., and Oliver, R. P. (2014). Sensitivity to three *Parastagonospora nodorum* necrotrophic effectors in current Australian wheat cultivars and the presence of further fungal effectors. *Crop Pasture Sci.* 65, 150–158. doi: 10.1071/CP13443

Tanaka, S. (1933). *Studies on black spot disease of the Japanese Pear (Pirus serotina Rehd). Memoirs of the College of Agriculture*. Kyoto: Kyoto University, 31.

Tomas, A., Gene, G. H., Reeck, G. R., Bockus, W. W., and Leach, J. E. (1990). Purification of a cultivar-specific toxin from *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. *Mol. Plant Microbe Interact.* 3, 221–224. doi: 10.1094/MPMI-3-221

Trujillo, M., Ichimura, K., Casais, C., and Shirasu, K. (2008). Negative regulation of PAMP-triggered immunity by an E3 ubiquitin ligase triplet in Arabidopsis. *Curr. Biol.* 18, 1396–1401. doi: 10.1016/j.cub.2008.07.085

Tuinstra, M. R., Ejeta, G., and Goldsbrough, P. B. (1997). Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. *Theor. Appl. Genet.* 95, 1005–1011. doi: 10.1007/s001220050654

Vincent, D. L., Du Fall, L. A., Livk, A., Mathesius, U., Lipscombe, R. J., Olliver, R. P., et al. (2012). A functional genomics approach to dissect the mode of action of the *Stagonospora nodorum* effector protein SnToxA in wheat. *Mol. Plant Pathol.* 13, 467–482. doi: 10.1111/j.1364-3703.2011.00763.x

VSN International (2011). *GenStat for Windows*, 14th Edn. Hemel Hempstead: VSN International.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

Waters, O. D. C., Lichtenzveig, J., Rybak, K., Friesen, T. L., and Oliver, R. P. (2011). Prevalence and importance of sensitivity to the *Stagonospora nodorum* necrotrophic effector SnTox3 in current Western Australian wheat cultivars. *Crop Pasture Sci.* 62, 556–562. doi: 10.1071/CP11004

Xia, Y. (2004). Proteases in pathogenesis and plant defence. *Cell. Microbol.* 6, 905–913. doi: 10.1111/j.1462-5822.2004.00438.x

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zeng, L. R., Qu, S., Bordeos, A., Yang, C., Baraoidan, M., Yan, H., et al. (2004). *Spotted leaf11*, a negative regulator of plant cell death and defense, encodes a U-box/armadillo repeat protein endowed with E3 ubiquitin ligase activity. *Plant Cell* 16, 2795–2808. doi: 10.1105/tpc.104.025171

Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi: 10.1038/ng.546

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AT and handling Editor declared their shared affiliation.

Theoretical and Applied Genetics
https://doi.org/10.1007/s00122-019-03517-8

ORIGINAL ARTICLE

Genetic analysis of wheat sensitivity to the ToxB fungal effector from *Pyrenophora tritici-repentis*, the causal agent of tan spot

Beatrice Corsi[1] · Lawrence Percival-Alwyn[1] · Rowena C. Downie[1,2] · Luca Venturini[3] · Elyce M. Iagallo[4] · Camila Campos Mantello[1,6] · Charlie McCormick-Barnes[1,5] · Pao Theen See[4] · Richard P. Oliver[4] · Caroline S. Moffat[4] · James Cockram[1]

Received: 14 August 2019 / Accepted: 17 December 2019

Abstract

Key message **Genetic mapping of sensitivity to the *Pyrenophora tritici-repentis* effector ToxB allowed development of a diagnostic genetic marker, and investigation of wheat pedigrees allowed transmission of sensitive alleles to be tracked.**

Abstract Tan spot, caused by the necrotrophic fungal pathogen *Pyrenophora tritici-repentis*, is a major disease of wheat (*Triticum aestivum*). Secretion of the *P. tritici-repentis* effector ToxB is thought to play a part in mediating infection, causing chlorosis of plant tissue. Here, genetic analysis using an association mapping panel ($n = 480$) and a multiparent advanced generation intercross (MAGIC) population (n founders = 8, n progeny = 643) genotyped with a 90,000 feature single nucleotide polymorphism (SNP) array found ToxB sensitivity to be highly heritable ($h^2 \geq 0.9$), controlled predominantly by the *Tsc2* locus on chromosome 2B. Genetic mapping of *Tsc2* delineated a 1921-kb interval containing 104 genes in the reference genome of ToxB-insensitive variety 'Chinese Spring'. This allowed development of a co-dominant genetic marker for *Tsc2* allelic state, diagnostic for ToxB sensitivity in the association mapping panel. Phenotypic and genotypic analysis in a panel of wheat varieties post-dated the association mapping panel further supported the diagnostic nature of the marker. Combining ToxB phenotype and genotypic data with wheat pedigree datasets allowed historic sources of ToxB sensitivity to be tracked, finding the variety 'Maris Dove' to likely be the historic source of sensitive *Tsc2* alleles in the wheat germplasm surveyed. Exploration of the *Tsc2* region gene space in the ToxB-sensitive line 'Synthetic W7984' identified candidate genes for future investigation. Additionally, a minor ToxB sensitivity QTL was identified on chromosome 2A. The resources presented here will be of immediate use for marker-assisted selection for ToxB insensitivity and the development of germplasm with additional genetic recombination within the *Tsc2* region.

Communicated by Kevin Smith.

Caroline S. Moffat and James Cockram have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-019-03517-8) contains supplementary material, which is available to authorized users.

✉ Caroline S. Moffat
caroline.moffat@curtin.edu.au

✉ James Cockram
james.cockram@niab.com

[1] John Bingham Laboratory, NIAB, Huntingdon Road, Cambridge CB3 0LE, UK

[2] Plant Sciences Department, University of Cambridge, Cambridge, UK

[3] Life Sciences Department, Natural History Museum, Cromwell Road, London SW7 5BD, UK

[4] Centre for Crop and Disease Management, School of Molecular and Life Sciences, Curtin University, Perth, Australia

[5] Faculty of Biology, Medicine and Health, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

[6] Present Address: Genetracer Biotech, Calle Albert Einstein 22, 39011 Santander, Spain

Introduction

Tan spot, also known as yellow (leaf) spot, is a major fungal disease of wheat (*Triticum aestivum* L.). It is caused by the necrotrophic fungus *Pyrenophora tritici-repentis* (Died.) Drechs. (abbreviated here as *Ptr*; syn,

Published online: 08 January 2020

Dreschlera tritici-repentis [*Dtr*]) and typified by necrotic lesions as well as regions of chlorosis on infected leaves, which result in reduced leaf photosynthetic area. This typically results in a 5–10% reduction in grain yield, although losses can reach 50% under favourable conditions (De Wolf et al. 1998). Tan spot is recognised as a major disease in a number of wheat-growing areas, including Europe, South America, Canada and Australia (Annone 1998; Ciuffetti et al. 2014; Savary et al. 2019). A complex *Ptr* race structure has been defined, with screening against a differential set of wheat varieties allowing at least eight races to be described, termed race 1–8 (Lamari et al. 2003). Necrotrophic effectors (previously termed 'host-selective toxins') mediate the interaction between a given *Ptr* race and its susceptible differential host line. In contrast to the classical gene-for-gene model, whereby interaction of avirulence effectors with host resistance gene complexes confers resistance, the tan spot host–pathogen system appears to be largely governed in an inverse gene-for-gene manner, whereby effector sensitivity is conferred by a single dominant host gene (Tan et al. 2010). The host genotype specificity of effectors makes them important factors in disease development, and the removal of host sensitivity genes is a priority for breeding efforts.

To date, three *Ptr* effectors have been described and demonstrated to be pathogenicity factors: ToxA, ToxB and ToxC. Each *Ptr* race is largely differentiated by its expression of one or a combination of these three effectors (Lamari et al. 2003). ToxA was the first *Ptr* effector to be isolated and is the most well studied (Ballance et al. 1989; Tomas et al. 1990; Faris et al. 1996). The majority of *Ptr* isolates worldwide produce ToxA (Friesen et al. 2005) comprising of races 1, 2, 7 and 8. ToxA triggers necrosis in wheat lines carrying susceptible alleles at the *Tsn1* locus. The mature PtrToxA protein encodes a 13.2 kDa peptide containing a fibronectin type III-like domain (Balance et al. 1996; Ciuffetti et al. 1997) and includes an arginyl-glycyl-aspartic acid (RGD) motif thought to be important for receptor binding and internalisation into wheat cells (Meinhardt et al. 2002; Manning et al. 2008). The *ToxA* gene is thought to have been transferred to *Ptr* via horizontal gene transfer from another wheat necrotrophic pathogen, *Parastagonospora nodorum* (Friesen et al. 2006). In wheat, the *Tsn1* locus confers sensitivity to ToxA from both *Ptr* and *P. nodorum*. *Tsn1* encodes a predicted protein containing a nucleotide-binding site leucine-rich repeat (NBS-LRR) domain and a serine/threonine protein kinase (S/TPK) domain (Faris et al. 2010). Wheat varieties insensitive to ToxA predominantly carry a complete deletion of *Tsn1*. However, while *Tsn1* is necessary to mediate ToxA recognition, several yeast-2-hybrid studies have reported different interacting proteins in the host. Early work showed that ToxA interacts with plastocyanin (Tai et al. 2007) and the chloroplast ToxABP1 protein (Manning et al. 2007), both of which may promote the induction of reactive oxygen species leading to cell death. More recently ToxA has been demonstrated to interact with the wheat pathogencitiy-related PR-1-5 protein in a highly specific manner (Lu et al. 2014). ToxA-induced necrosis was enhanced by co-infiltration of both ToxA and PR-1-5 into sensitive wheat lines, but not with a non-interacting PR-1-5^{N141A} mutant (Lu et al. 2014). Therefore, the interaction of ToxA with the PR-1-5 protein appears to play a role in promoting necrosis in *Tsn1*-containing wheat. The role of ToxA as a key tan spot disease determinant has been demonstrated in interactions with many but not all wheat cultivars (Ciuffetti et al. 1997). Recent work screening Australian wheat lines with a *Ptr* mutant carrying a deletion of *ToxA* has further highlighted the importance of ToxA in tan spot disease (See et al. 2018). Disease levels were significantly reduced on about one third of the *Tsn1* wheat lines. The observation that disease levels did not always decrease highlights that while ToxA is a key determinant in tan spot disease, it is not the whole story.

Compared to the ToxA-*Tsn1* interaction which results in rapid necrosis, the ToxB-*Tsc2* and ToxC-*Tsc1* interactions both result in slower chlorotic responses in sensitive wheat lines. ToxC is produced by *Ptr* races 1, 3, 6 and 8 (Strelkov and Lamari 2003). Partial culture filtrate purification based on gel filtration, ion exchange and reverse-phase chromatography indicated that ToxC is a polar, non-ionic, low molecular mass molecule (Effertz et al. 2002). ToxC-induced chlorosis was observed on wheat cultivars carrying sensitive alleles at the *Tsc1* locus on the short arm of chromosome 1A (Effertz et al. 2002).

ToxB is a small protein that causes necrosis in sensitive wheat lines. It was first identified in the culture filtrate of race 5 *Ptr* isolates, as well as in combination with other effectors in races 6, 7 and 8 (Strelkov et al. 2002; Lamari et al. 2003). The *ToxB* gene encodes a mature protein of 64 amino acids (Martinez et al. 2001), although no functional motifs that might contribute to toxic activity have been identified to date (Ciuffetti et al. 2010). *ToxB* copy number variation has been shown to impact on tan spot susceptibility, with isolates containing increased *ToxB* copies linked to higher *ToxB* gene expression and protein production and the induction of increased symptoms in the host (Strelkov et al. 2002; Strelkov and Lamari 2003; Martinez et al. 2004). ToxB-mediated chlorosis is known to be light dependent (Strelkov et al. 1998), and ToxB has been shown to inhibit photosynthesis and to modify the wheat leaf proteome prior to the development of chlorosis (Kim et al. 2010). The major ToxB sensitivity locus *Tsc2* has previously been mapped to the short arm of chromosome 2B (Friesen and Faris 2004; Abeysekara et al. 2010), accounting for up to 69% of total phenotypic variation (Friesen and Faris 2004).

Although ToxB was first characterised more than 17 years ago (Martinez et al. 2001), relatively little systematic

investigation of wheat variety sensitivities has been undertaken to date, and the relationship between ToxB sensitivity and disease susceptibility has not been well defined. In this study, we present information on ToxB sensitivity in a collection of over 470 European wheat varieties and undertake genetic mapping of ToxB sensitivity using a combination of an association mapping panel, and an eight parent multiparent advanced generation intercross (MAGIC) population. Based on genotypic data generated using a 90 k single nucleotide polymorphism (SNP) array, fine mapping of the ToxB sensitivity locus *Tsc2* allowed development of a closely linked SNP-based marker for wheat disease resistance research and breeding.

Methods

Wheat germplasm and genotypic data

Three sets of wheat germplasm were used to assess ToxB sensitivity and for genetic mapping. (1) Parents of wheat genetic mapping populations: the eight founders of the 'NIAB Elite MAGIC' population (Mackay et al. 2014), eight founders of the 'BMW MAGIC' population (Stadlmeier et al. 2018), and 14 additional founders of various bi-parental populations and key germplasm stocks from hexaploid and tetraploid wheat (Table 1). (2) A wheat association mapping (AM) panel consisting of 480 predominantly UK, French and German varieties drawn from historic collections and National Lists, encompassing varieties released between 1916 and 2007 (Downie et al. 2018; Supplementary Table 1). (3) The 'NIAB Elite MAGIC' population: eight founders and 643 progeny. The AM and 'NIAB Elite MAGIC' populations were previously genotyped using an Illumina iSelect 90,000 feature single nucleotide polymorphism (SNP) wheat array (Wang et al. 2014). The AM panel data matrix consisted of 22,237 polymorphic SNPs (minor allele frequency $\geq 6\%$) across 480 varieties (data available at https://www.niab.com/pages/id/326/Resources), while the MAGIC data matrix consisted of 20,643 polymorphic SNPs across 643 progeny (Mackay et al. 2014; Gardner et al. 2016).

ToxB effector production

ToxB protein was heterologously expressed via an *Escherichia coli* SHuffle strain and purified using immobilised metal affinity chromatography (IMAC) as described in See et al. (2019). Briefly, the DNA sequence that encodes the mature ToxB protein (GenBank accession PZD27634) (Martinez et al. 2001) was cloned into the pET21a(+) (Novagen) vector using the primers ToxB-sigP_F (5′-GGAATTCCATATG AACTGCGTCGCCAATAT-3′) and ToxB_R (5′-CCGCTC GAGACAACGTCCTCCACTTTGCA-3′) that included the engineered *NdeI* and *XhoI* sites (underlined), respectively. The resulting construct contained the predicted ToxB protein fused to a poly-histidine tag at the C-terminal, and was used to transform the *E. coli* SHuffle strain. The SHuffle strain harbouring the ToxB expression construct was grown in Terrific Broth medium and induced with β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 100 µM for the expression of ToxB. Cells were harvested by centrifugation and resuspended in 5 ml binding buffer (20 mM sodium phosphate, 40 mM imidazole, 500 mM NaCl, pH 7.4) for the purpose of downstream his-tag purification. Cells were then lysed using sonication of 10-s pulses with 10-s intervals for 1 min (Sonoplus HD 3100, Bandelin, Germany). The cell debris was pelleted by centrifugation, and the ToxB protein was purified from the cell extract using HisPur Ni-NTA purification spin columns (Thermo Scientific) by gravity flow according to manufacturer's instructions. Purified protein was dialysed in 20 mM sodium phosphate buffer, pH 7.4, quantified via the bicinchoninic acid assay (Smith et al. 1985) and stored in freeze-dried form at − 80 °C. Prior to infiltration, protein was resuspended in sterile water to a concentration of 200 µg/ml.

Experimental design

All germplasm was sown into 96-well trays filled with M2 compost (Levington, Everiss). The experimental plan for each germplasm set comprised five replicates of each variety or accession, arrayed in a randomised block design using R/blocksdesign (cran.r-project.org). Seeds of the first germplasm set were grown in a growth chamber (Conviron) using a 16 h light (at 20 °C) and 8 h dark (15 °C) photoperiod. The AM and MAGIC populations were grown in a heated glasshouse (16 h light at 20 °C, 8 h dark at 17 °C) for 14 days with supplementary lighting to maintain photoperiod. ToxB infiltration was undertaken 14 days after sowing, following the method described by Moffat et al. (2014). Briefly, the first leaf of seedlings at growth stage 12 (GS12, Zadoks et al. 1974) was infiltrated with 50 µl of ToxB suspension at a concentration of 200 µg/mL, and the extent of infiltration along the leaf marked using a non-toxic pen. Phenotyping of ToxB sensitivity was carried out a week after infiltration and scored using a 0–5 scale, as described by See et al. (2019). A score of 0 = no visible symptoms; 1 = slight chlorosis; 2 = full chlorosis; 3 = extensive chlorosis with/without slight necrosis; 4 = chlorosis with necrosis; 5 = full necrosis. A water control was also used to establish a symptom baseline for the evaluation of possible damage during the infiltration process.

Table 1 Mean ToxB sensitivities for 29 wheat varieties

Variety	Populations and resources	Mean ToxB sensitivity
Alchemy	NIAB Elite MAGIC[1]	0.1
Ambition	BMW MAGIC[2]	0
Apogee	Apogee × Paragon[3]	0
Avalon	Avalon × Cadenza[4]. Avalon × Cadenza NILs[5]	0
Batallion	Oakley × Batallion[6]	2
BAYP4535	BMW MAGIC[2]	0
Brompton	NIAB Elite MAGIC[1]	0
Bussard	BMW MAGIC[2]	0
Cadenza	Avalon × Cadenza[4]. Avalon × Cadenza NILs[5]. TILLING[7]	0.2
CS	Paragon × CS[5]. Reference genome[8]	0
Claire	NIAB Elite MAGIC[1]. Claire × Malacca[9]. Genome sequence[10]	0.2
Dic12b[†]	Tios × Dic12b[11]	0
Event	BMW MAGIC[2]	0
Exsept	Oakley × Exsept[6]	0
Firl3565	BMW MAGIC[2]	0
Format	BMW MAGIC[2]	0
Gatsby	Oakley × Gatsby[6]	0
Hereward	NIAB Elite MAGIC[1]	0.2
Julius	BMW MAGIC[2]	0
Malacca	Claire × Malacca[9]	0
Oakley	Oakley × Batallion[6], Oakley × Exsept[6], Oakley × Gatsby[6]	0.1
Paragon	Paragon × CS[5], Paragon × SHW CSSL[12], genome sequence[10]	0
Potenzial	BMW MAGIC[2]	0
Rialto	NIAB Elite MAGIC[1]	0.2
Robigus	NIAB Elite MAGIC[1]. Genome sequence[10]	0
SHW-041	Paragon × SHW CSSLs[11]	0
Soissons	NIAB Elite MAGIC[1]	0.7
Tios[†]	Tios × Dic12b[11]	0
Xi19	NIAB Elite MAGIC[1]. Solstice × Xi19[9]	2.8

Sensitivity was scored on a 0 (no sensitivity) to 4 (highly necrotic) scale (See et al. 2019), with five replicates per genotype. The relevant major resources associated with each variety are listed. CS = cv. 'Chinese Spring'. All varieties are hexaploid (*T. aestivum*), apart from those accessions indicated: [†]tetraploid *T. turgidum* subsp. *dicoccum*. Resources listed are bi-parental populations, unless otherwise indicated. CSSL = chromosome segment substitution lines. NILs = near isogenic lines. References: 1 = Mackay et al. 2014. 2 = Stadlmeir et al. 2018. 3 = Allen et al. 2017. 4 = developed by C. Ellerbrook, L. Sayers, and T. Worland (John Innes Centre, UK). 5 = developed within the Wheat Genetic Improvement Network (WGIN) project, https://www.wgin.org.uk/. 6 = ERYCC report, available at https://cereals.ahdb.org.uk/media/200035/pr496.pdf. 7 = Krasileva et al. 2017. 8 = IWGSC RefSeq v1.0, available at https://wheat-urgi.versailles.inra.fr/Seq-Repository/Assemblies. 9 = Breeder's population. 10 = available at https://wheatis.tgac.ac.uk/grassroots-portal/blast. 11 = developed within the Biotechnology and Biological Sciences Research Council (BBSRC) project, 'WISP' (BBSRC grant reference BB/I002561/1)

Statistical analyses and QTL mapping

Summary statistics (mean, median, standard deviation and variance) were calculated using the software GenStat (VSN International, 16th edition). Best linear unbiased estimates (BLUEs) were calculated using a linear mixed approach in REML using GenStat (VSN International 2015). Heritability was calculated using GenStat: Broad sense heritability of line means was calculated from the estimate of the variance components in REML, taking into account all features of the experimental designs. Heritability was then estimated as $h^2 = \sigma^2 G / (\sigma^2 G + \sigma^2 e)$ where $\sigma^2 G$ is the genetic variance of line means and $\sigma^2 e$ is the residual variance.

Genome-wide association scans (GWAS) using the AM panel were undertaken using the Efficient Mixed-Model Association algorithm (Kang et al. 2008) using a compressed mixed linear model that includes both fixed and random effects (Zhang et al. 2010), implemented with the Genome

Association and Prediction Integrated Tool (GAPIT) package (Lipka et al. 2012) in R (R Core Team 2013). Genetic stratification in the population was corrected for using a kinship matrix constructed in GAPIT derived from a subset of SNPs skimmed from the complete set using an r^2 threshold > 0.75. Further GWAS analysis was undertaken using selected markers as co-factors in the analysis, following the guidelines in the GAPIT manual. A Bonferroni-adjusted $P = 0.01$ ($-\log_{10}P = 6.35$) significance threshold was used.

MAGIC QTL mapping was carried out using 7369 unique mapped SNPs, with genetic map positions as described by Gardner et al. (2016). Four QTL analysis approaches were used: (1) SMA (single marker analysis): regression against single markers using R/lme4. (2) IBD (identity by descent): regression against haplotype probability estimates calculated using the 'mpprob' function in R/mpMap (Huang and George 2011) implemented in R/qtl (Broman et al. 2003) with a threshold of 0.5. (3) IM (interval mapping): conducted in R/mpMap using R/mpMap haplotype probability estimates. (4) CIM (composite interval mapping): conducted in R/mpMap with 5 or 10 covariates using R/mpMap haplotype probability estimates. For IBS and IBD analyses, multiple-test correction was carried out using *R/q* value, with a threshold of $q < 0.05$. For IM/CIM an empirical *P* threshold of 0.05 was determined for QTL analyses by conducting 100 simulations, using the sim.sigthr function in *R*. This value, together with a window size of 100 markers was used to determine QTL peaks using 'find.qtl'. 'Fit.qtl' was then applied, and QTL retained which had $P < 0.05$ in the fitted model, as well as percentage variation explained > 1%.

SNP anchoring and pedigree analyses

Selected SNPs were anchored to the wheat cv. Chinese Spring 42 IWGSC RefSeq v1.0 physical map (IWGSC 2018) by BLASTn (Altschul et al. 1990). Where BLASTn hits of equal match were identified on multiple chromosomes, genetic map position (Gardner et al. 2016) was used to assign hits to chromosomes. Wheat pedigree information was obtained from Fragley et al. (2019), with the underlying data available at https://www.niab.com/pages/id/326/Resources). The pedigree was displayed using Helium v.1.17.08.14 (Shaw et al. 2014) and images prepared using CorelDRAW (Corel Corporation, Canada).

KASP genotyping

To validate the conversion of selected SNPs from the Illumina 90 k array to the Kompetitive Allele-Specific PCR (KASP) genotyping platform (LGC Genomics, UK), DNAs were extracted from the eight MAGIC founder varieties. Marker co-dominance was investigated using a 50:50 mix of DNAs from two varieties known from the 90 k SNP dataset to contrast for allele call. Additionally, DNA was extracted from a panel of 48 UK varieties that post-date the AM panel, released to the AHDB Recommended List between 2009 and 2017 (Supplementary Table 2). All DNAs were extracted from two week old leaves using a modified Tanksley protocol (Fulton et al. 1995), and concentration was determined using a Nanodrop 200 spectrophotometer (Thermo Scientific). DNA sequences flanking each targeted SNP were used to design KASP primers using the software PolyMarker (Ramirez-Gonzalez et al. 2015). Primers were synthesised by Sigma-Aldrich (Cambridge, UK) and KASP genotyping undertaken as described by Cockram et al. (2015). The results were visualised using SNP Viewer v.1.99 (https://lgcgenomics.com/). The subset of plants that post-dated the AM panel were grown and phenotyped for ToxB sensitivity and genotyped with the KASP marker, following the methods described above.

Genomics analyses of the Tsc2 region of synthetic wheat line W7984

Analysis of the *Tsc2* region was carried out in eleven steps. (1) The W7984 assembly scaffolds (Chapman et al. 2015) were aligned to the reference 'Chinese Spring' genome, RefSeq v1.0 (IWGSC 2018) with MiniMap2 v2.15 (Li 2018) using command line options: '-I 20G -ax asm5 –secondary = no'. Scaffolds with primary alignments to a region conservatively encompassing the *Tsc2* locus (2B: 21536610..0.27030113 bp) (SAM flags 0 and 16) were extracted from the W7984 assembly and ordered and orientated based on their SAM file coordinates and mapping orientation flags before being joined together as one fragment with each contributing scaffold separated by 100 Ns. (2) The synthetic 2B assembled fragment and the corresponding region of CS42 (chr2B: 21536610..0.27030113 bp) were then repeat masked using Repeatmasker v4.0.8 (https://www.repeatmasker.org) using the command line option: '-species Triticum aestivum'. (3) The two repeat masked sequences were aligned using minimap2 v2.16, using the 'Chinese Spring' region as reference and with the command line options:'–cs = long -c -t 20 -K 500 M –Y'. Methods for the liftover of the 'Chinese Spring' gene models are further described in Supplementary Text 1. (4) RNA-seq reads (> 8 billion) from BioProject accession numbers PRJDB2496, PRJEB12497, PRJEB25593, PRJEB25639, PRJEB25640, PRJNA213168, PRJNA243835 were obtained from the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA). Further details on the processing of the RNA-seq data and subsequent transcript assembly, gene model and splice junction prediction are listed in Supplementary Text 1. (5) Protein datasets sourced from the genome assemblies of three cereal species (listed in Supplementary Text 1) were aligned to our synthetic 2B fragment from 'W7984' that spanned the *Tsc2* region using GenomeThreader v1.7.1 (Gremme et al. 2005). Protein alignments

were further filtered using the script 'filter_exonerate.py' from the ei-annotation suite (https://github.com/lucventurini/ei-annotation), using the Portcullis junctions above as externally validated junctions and with further parameters: '-minI 20 -maxE 1000 -maxM 5000', which excluded any hit with terminal introns longer than 1 kbps or internal introns longer than 5 kbps. (6) Transcript reconstruction: Mikado (Venturini et al. 2018) leverages transcript assemblies generated by multiple methods to improve transcript reconstruction. Loci were first defined across all input assemblies with each assembled transcript scored based on metrics relating to open reading frame (ORF) and cDNA size, relative position of the ORF within the transcript, un-transcribed region (UTR) length and presence of multiple ORFs. Mikado was used to integrate the Illumina assemblies generated above, as described in further detail in Supplementary Text 1. (7) Gene predictions: the information obtained from steps 1–6 was used for final gene prediction using Augustus (Stanke and Morgenstern 2005) as described in Supplementary Text 1. (8) Determining the likelihood of gene presence/absence in 'Chinese Spring': Percentage gene presence in 'Chinese Spring' (CS) was determined by using BLAST to query the CS chr2B region for each CDS followed by calculating the mean average of the best % CDS presence (nident/qlen * 100) for each transcript. Where there was more than one splice variant present, the transcript with the highest average CDS presence was used to represent the % gene presence in CS. (9) Functional annotation of the synthetic 2B assembled fragment predicted genes was carried out via BLAST queries to the uniprot90 protein database (releaseDate = '2019-04-10' version = '2019_03'). For each predicted peptide, the top 20 BLAST hits (based on e-values ≤ 1e−25) were retained. The gene functions were then assigned by hand. Where there was no function known or where BLAST results were > 1e−25, proteins were labelled as uncharacterised. (10) Identifying the corresponding CS gene IDs to the synthetic gene predictions: The coding sequences for each of the transcripts from both the synthetic region and corresponding CS region were extracted using gffread. BLAST was used to query the CS coding sequences with the predicted synthetic coding sequences. The best BLAST results (determined by e-value and bit score) were taken for each query and coding sequences with less than a 90% CDS identity (nident / qlen * 100) were ignored. (11) Circos (Krzywinski et al. 2009) plots were constructed using a 10-kb jumping window for both gene and repeat counts. Marker positions within the synthetic assembled region were determined using BLAST. Canonical kmer repeats (> 1) are 91 bp and contain no *N* labelled bases. Candidate genes were identified using the following search terms from the Uniprot functional annotation 'resistance', 'defence' 'disease', 'wall-associated', 'lrr', 'nbs', 'stress' and 'cysteine-rich'.

Results

ToxB sensitivity in a panel of wheat genetic mapping population founders

ToxB infiltration of the panel of 29 parental lines of wheat mapping populations found moderate to high sensitivity scores (≥ 2) in just two accessions (Table 1). All eight founders of the 'BMW MAGIC' population were insensitive (score < 0.2) to ToxB, ruling out this population for QTL analysis. Of the eight 'NIAB Elite MAGIC' founders, six lines were insensitive (score ≤ 0.2; 'Alchemy', 'Brompton', 'Claire', 'Hereward', 'Rialto', 'Robigus') one showed intermediate sensitivity ('Soissons', score = 0.7), while one was highly sensitive ('Xi19', score = 2.8). Among the remaining 12 varieties tested, only 'Battalion' showed marked ToxB sensitivity (score = 2), with all other lines classified as ToxB insensitive (score ≤ 0.2; 'Cadenza', 'Oakley', 'Apogee', 'Avalon', 'Chinese Spring', 'Dic12b', 'Exsept', 'Gatsby', 'Malacca', 'Paragon', 'SHW-041' and 'Tios'). Based on these results, two wheat populations were selected for further analysis: 'NIAB Elite MAGIC', and a European AM panel so as to screen as widely as practicable for the evidently rare cases of sensitivity.

Genetic analysis of ToxB sensitivity in the MAGIC and AM populations

ToxB sensitivity in the MAGIC population ($n = 643$) was found to be highly heritable ($h^2 = 0.95$). Mean ToxB sensitivity in the progeny ranged between 0 (insensitive) and 2.25, with 503 lines = 0, 70 lines > 0 ≤ 1 and 67 lines > 1. QTL mapping of ToxB sensitivity in the MAGIC population using $\log_{10}$ transformed phenotypic data and four genetic mapping approaches (SMA, IBD, IM, CIM) all identified a major QTL on the short arm of chromosome 2B (e.g. for IM analysis, $-\log_{10}P = 109.19$, accounting for 41% of the phenotypic variation) (Fig. 1; Table 2; Supplementary Table 1). To help determine the *Tsc2* interval, CIM using 5 and 10 covariates and an empirical $P = 0.01$ significance threshold ($-\log_{10}P = 5.29$) identified a 15.11 cM region (from SNPs BS00034887_51 to BS00053520_51). Within this region, via systematic comparison of the *P* values along with the genetic and physical map positions of the markers, we define the peak of this QTL to be located in a 1.094 Mbp interval from 24.092 Mbp (BS00070051_51) to 25.186 Mbp (BS00002660_51), which includes all markers with $-\log_{10}P > 90$ from the IM and CIM analyses. This locus corresponds to be the major ToxB sensitivity locus *Tsc2* (Friesen and Farris 2004), based on the position of the MAGIC SNPs identified here

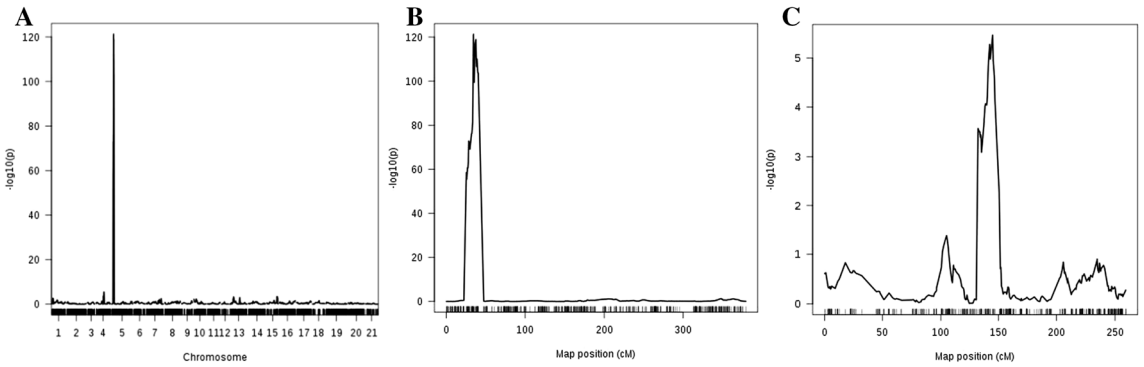


Fig. 1 An example of the results of genetic analysis of ToxB sensitivity in the 'NIAB Elite MAGIC' population. Presented here are the outputs using composite interval mapping (CIM), covariates = 10. (A) QTLs found across all 21 wheat chromosomes with 1 (1A), 2 (1B), 3 (1D), 4 (2A) through to 21 (7D). (B) Chromosome 2B, showing the major ToxB sensitivity locus, *Tsc2*. (C) Chromosome 2A, showing the putative minor ToxB sensitivity QTL located at ~ 142 cM

Table 2 Summary of the QTL identified for ToxB sensitivity using the 'NIAB Elite MAGIC' population

Chr	Left marker name (IWGSC RefSeq v1.0 Mb position)	Right marker name (IWGSC RefSeq v1.0 Mb position)	$-\log_{10}P$	PVE (%)
2B	BS00070050_51 (24.09)	Kukri_c63748_1453 (25.02)	109.19	41.0
2A	CAP12_rep_c7918_56 (613.00)	RAC875_c38018_278 (639.99)	5.70	1.4

The chromosome 2B QTL, *Tsc2*, was identified with all mapping methods tested: IM (data presented in the table above) and CIM using 5 and 10 covariates, IBS and IBD, with the best QTL resolution obtained with CIM-cov5 and -cov10. The 2A QTL was identified using CIM-cov10 (data presented in the table) but was not significant at CIM-cov5 ($-\log_{10}P = 3.68$) using the empirical $P = 0.05$ threshold. Flanking markers listed represent those immediately adjacent to the peak of the QTL, based on genetic map order. Chr = chromosome. PVE = phenotypic variation explained. Further details for the markers with all QTL intervals are listed in Supplementary Table 1

and the previously identified *Tsc2*-linked flanking markers *TC339813* and *BE444541* (Abeysekara et al. 2010), which we find to be located on chromosome 2B at 22.230 and 24.100 Mb, respectively. Additionally, a putative minor QTL was identified on the long arm of chromosome 2A, identified using CIM-cov10 ($-\log_{10}P = 5.70$) (Fig. 1; Table 2). This QTL is located between 132.174 and 146.829 cM with the QTL peak defined by SNP markers Kukri_c24064_2095 and BS00022641_51 and explained 1.4% of the phenotypic variation. Given the locations of the 2A and 2B QTLs on the long and short chromosomal arms, respectively, these are not assumed to be homoeologous loci. Note that for the minor 2A QTL, while the CIM-cov10 result was above the empirical $P = 0.05$ ($-\log_{10}P = 4.36$) and $P = 0.01$ ($-\log_{10}P = 5.29$) significance thresholds, using CIM-cov5 it was found to be insignificant ($-\log_{10}P = 3.68$).

ToxB sensitivity was also found to be highly heritable in the AM panel ($n = 480$, $h^2 = 0.9$). The majority of varieties (93%) were found to either be insensitive (ToxB sensitivity < 0.2) or slightly sensitive (sensitivity > 0.2, < 1), while just 7% of varieties were classified as sensitive (≥ 1) (Supplementary Table 2). GWAS using the 22,237 SNPs after skimming the complete data matrix on minor allele frequency ≤ 0.06 identified 73 significant SNPs at the Bonferroni corrected $P = 0.01$ threshold ($-\log_{10}P > 6.35$) (Fig. 2; Supplementary Table 3). Of these SNPs, 64 were located on the MAGIC genetic map on chromosome 2B between 15.13 and 73.65 cM. Anchoring the remaining nine SNPs to the physical map via BLASTn indicated eight mapped to chromosome 2B (Supplementary Table 3) and supported by the analyses of Gardner et al. (2016) who treated these SNPs as a trait and QTL mapped them back to chromosome 2B on the 'NIAB Elite MAGIC' genetic map. For the remaining unmapped SNP (Kukri_c30668_294), while BLASTn predicted it to map to 2D, using the SNP as a trait predicted it to be located on 2B (Gardner et al. 2016). Accordingly, we tentatively assign this marker to chromosome 2B.

Investigation of the haplotypes defined by the 12 SNPs genotyped across the peak of the *Tsc2* locus in the AM panel allowed a condensed five-SNP haplotype across the region to be determined (Supplementary Table 4). This defined 10 haplotypes, six of which were associated with low ToxB sensitivity (the 'group 1' haplotypes: 1.1 to 1.6) and four with high sensitivity (the 'group 2' haplotypes: 2.1 to 2.4), with the differences in ToxB sensitivity between

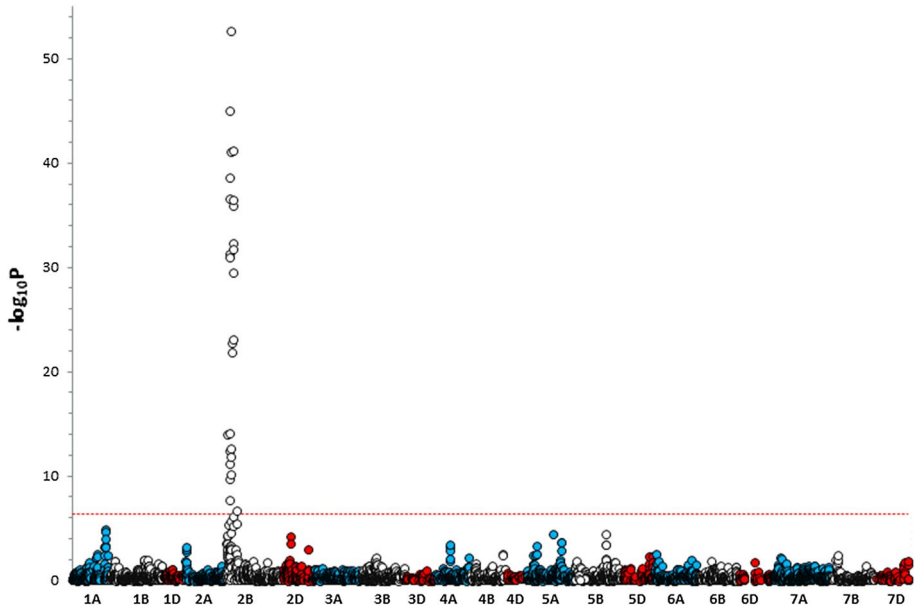


Fig. 2 Manhattan plot of GWAS results for ToxB sensitivity in the association mapping panel ($n = 480$) using 22,237 SNPs. Markers are ordered according to the 'NIAB Elite MAGIC' genetic map (Gardner et al. 2016). The Bonferroni-adjusted $P = 0.01$ significance threshold ($-\log_{10}P > 6.35$) is indicated. The 21 wheat chromosomes, 1A through to 7D, are indicated. Unmapped markers are not shown

Table 3 Comparison of allele scores at the *Tsc2*-linked SNP marker BS00072620_51 with the number of varieties in the AM panel found to have low and high ToxB sensitivity (sensitivity scores < 1 and ≥ 1 ≤ 3, respectively)

BS00072620_51 genotype	ToxB sensitivity < 1	ToxB sensitivity ≥ 1	Totals
A:A	442	0	442
B:B	0	38	38
	442	38	480

The 480 accessions used represent those with both genotype and phenotype data

these two haplotype groups found to be highly significant (Welch two sample *t*-test: $t = 25.5$, $P \sim 0$) (Supplementary Figure 1). Additionally, clear recombinations proximal (haplotype 2.2) and distal (haplotypes 2.3 and 2.4) to *Tsc2* delimited the target region to a 1921-kb region between 231.063 Mbp (TraesCS2B01G046400) and 250.180 Mbp (TraesCS2B01G051000) in the AM panel. All three of the most significant SNPs identified by GWAS ($-\log_{10}P = 52.63$; BS00070050_51, BS00072620_51 and BS00075303_51) originated from gene model TraesCS2B01G048500, and all three cosegregate on the 'NIAB Elite MAGIC' genetic map at 34.38 cM. Additionally, BS00070050_51 was also identified as the marker immediately proximal to the *Tsc2* QTL peak in the MAGIC IM and CIM analyses. In the AM panel, this SNP perfectly predicts sensitivity to ToxB, with alleles A:A and G:G diagnostic for sensitivity (score ≥ 1) and insensitivity (score ≤ 0.6), respectively (Table 3; Supplementary Table 2). When GWAS was repeated using the *Tsc2*-lined marker BS00070050_51 as a cofactor, no additional significant associations were found.

Bioinformatic analysis of the Tsc2 region

In summary, the genetic analyses defined a 1921-kb region defined in the AM panel (23.106 to 25.027 Mbp), extended proximally to 25.236 Mbp by the proximal end of the MAGIC peak region. This combined interval of 2129 kb contains 108 predicted genes in the ToxB-insensitive variety 'Chinese Spring' (Supplementary Table 5). Within the interval, the most significant SNPs from the AM panel and the MAGIC population were all located within gene model TraesCS2B01G048500, a glyoxylate reductase/hydroxypyruvate reductase located at 24.092 Mbp. Additional SNPs identified in the AM panel were contained within the interval: hit 6 (GENE-1343_556) originates from TraesCS2B01G048700 which encodes an arginase while hit 12 (Kukri_c63748_1453) is predicted to encode a cullin-associated NEDD8-dissociated protein 1. The wheat gene controlling sensitivity to the *Ptr* effector ToxA encodes a protein kinase NBS-LRR protein (Faris et al. 2010). A gene model predicted to encode an NBS-LRR protein is located within the 2129-kb 'Chinese Spring' region identified here (TraesCS2B01G050500). However, TraesCS2B01G050500 does not show high sequence similarity to *Tsn1*. Additionally, gene model TraesCS2B01G051200 annotated as an *MLO*-like gene, was located towards the distal end of the MAGIC QTL peak. Several additional disease resistance pathway genes were located just outside the *Tsc2* interval, including an *NBS-LRR* gene from which hits 9 and 13 originated (TraesCS2B01G045700), and surrounded by a wider cluster of seven additional NBS-LRR genes.

'Chinese Spring', for which the wheat reference genome is available, is insensitive to ToxB. Based on this, as well as the existing evidence in the published literature and the

identification of a single ToxB sensitivity locus via analysis of the association mapping panel undertaken here, Chinese Spring would be expected to contain a non-functional allele/deleted gene at the *Tsc2* locus. Of the few additional hexaploid wheat varieties for which a genome sequence is currently available, only one has previously been shown to be ToxB sensitive: 'W9784' (Friesen and Faris 2004). This is a synthetic hexaploid wheat, derived by the hybridisation of the tetraploid wheat species *Triticum durum* Desf. with the diploid wheat progenitor *Aegilops tauschii*. Although the contig length of the 'W9784' de novo genome assembly (Chapman et al. 2015) represents just 55% (7.88 Gb) of the 14.27 Gb 'Chinese Spring' RefSeq v1.0 assembly (IWGSC 2018), we nevertheless attempted to exploit this resource to investigate possible gene content in the *Tsc2* region. We first aligned the 'W9784' assembly to the 'Chinese Spring' genome, and then extracted primary alignments to a highly conservative physical interval across the 'Chinese Spring' *Tsc2* region (2B: 14,040,000–30,500,000 bp), spanning GWAS hit 15 (BS00085748_51, proximal) through to hit 23 (BS00022572_51, distal). This region comfortably includes the 2129-kb interval as defined from the combined AM and MAGIC analyses. The 'W9784' scaffolds were stitched together to form a single 'super-contig' (Fig. 3) consisting of 705 scaffolds with a mean scaffold size of 5.6 Kb. The ungapped assembled super-contig was 3.47 Mb representing 64.3% of the 'Chinese Spring' region (5.39 Mb). We then undertook gene annotation of the 'W9784' super-contig (see Supplementary Text 1), resulting in the identification of 379 gene models (Supplementary Table 6). After excluding gene

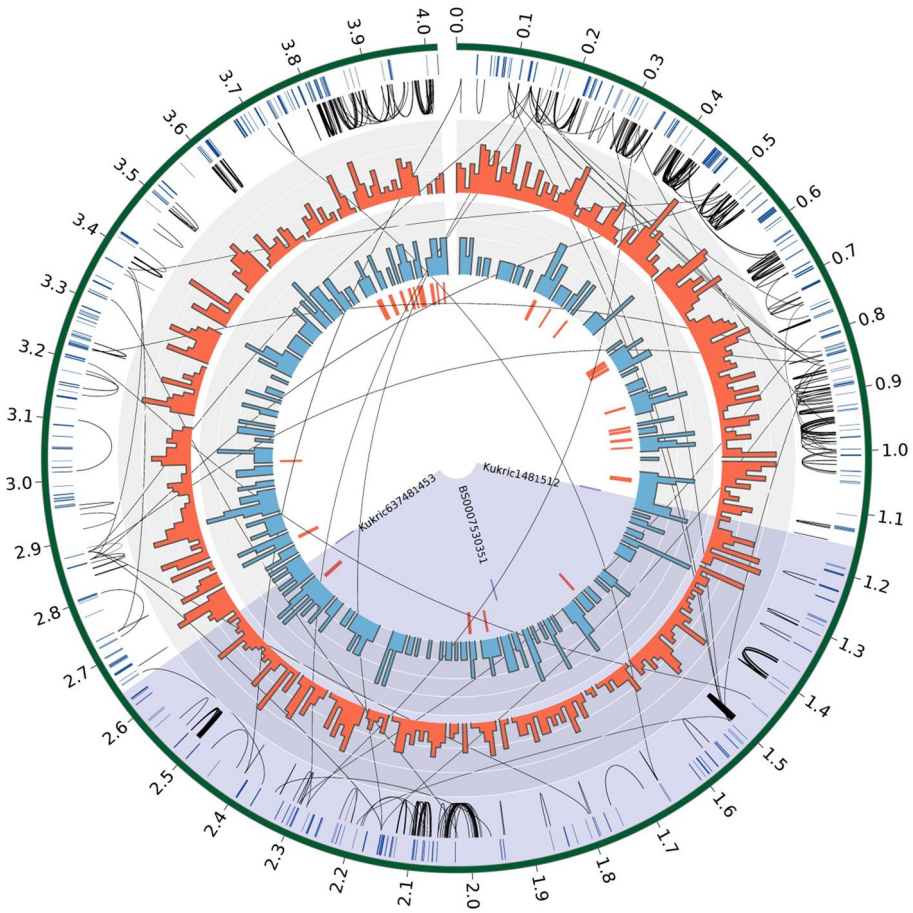


Fig. 3 Circos plot illustrating key genomic and genic features of the 'Synthetic W7984' 'super-contig' (705 assembly scaffolds, totalling 3.47 Mb) spanning the wider *Tsc2* locus, based on the 'Chinese Spring' region 2B: 14,040,000–30,500,000 bp. Tracks, from outside to inside: (1) the super-contig, with size indicated in Mb. (2) tick marks to illustrate start/end points between scaffolds, (3) loops to indicate sequence homology of genes based on 91-mers, (4) histogram of transposable element density, (5) histogram of gene density, (6) candidate genes, (7) tick marks indicating the named SNPs that delineate the boundaries of the most likely *Tsc2* region: (a) regions kukric1481512 to kukric637481453 (full names Kukri_c148_1512 and Kukri_c63748_1453, respectively) represent a conservative *Tsc2* interval, based on clear recombinations in the AM panel and encompassing the peak of the MAGIC QTL in this region, (b) region BS0007530351 (full name BS00075303_51, originating from the same RefSeqv1.1 gene model as BS00070050_51) to kukric637481453, which spans the peak of the MAGIC QTL and contains AM panel GWAS hits 1–6, 10 and 12

models annotated as transposable elements, 23 potential candidate genes were identified in the remaining 347 models, including four within the GWAS interval encompassing hits 1–8, 10 and 12 (from SNPs Kukri_c148_1512 to Kukri_c63748_1453; 'W9784' gene models jg131, jg166 and jg172) (Supplementary Table 6). Three of these candidates were annotated as LRR receptor-like serine/threonine protein kinases. Gene jg131 is collinear with 'Chinese Spring' gene model TraesCS2B01G047800, while jg166 is absent in the collinear region of 'Chinese Spring', and jg172 appears to be present (99% average CDS nucleotide identity), but not annotated as a gene in 'Chinese Spring'. In addition, a WRKY transcription factor (jg224), collinear with 'Chinese Spring' gene TraesCS2B01G050500 annotated as a TIR-NBS-LRR domain disease resistance protein, was identified. Of these, jg166, jg172 and jg224 were located within the interval defined by the peak of the MAGIC QTL (from SNPs Kukri_c148_1512 to Kukri_c63748_1453). Using a cut-off of 80% CDS sequence similarity, 51 genes in the 'W9784' assembly were found to be absent from the *Tsc2* region in 'Chinese Spring' (Supplementary Table 6). Details of all 'W9784' genes and candidate genes are summarised in Supplementary Table 6. As the contig size of the 'W9784' assembly is very low, with an N50 of 6.7 kb, we did not undertake further characterisation of its genic content. Nevertheless, this information provides a starting point with which to undertake future studies.

Development of a diagnostic KASP marker linked to Tsc2

The highly significant SNP BS00072620_51 identified in both the AM and MAGIC populations was selected for conversion from the 90 k SNP array to the KASP genotyping platform. The KASP primers designed (Allele-specific-A: 5′-tgattgcggagtatgtgca-3′, Allele-specific-B 5′- tgattgcggagtatgtgcg-3′, Common primer: 5′-gcaatgcgtgtccgtgtaaata-3′) were tested on DNAs extracted from all eight MAGIC founders, as well as a 50:50 mix of DNAs from 'Brompton and 'Xi19', and found tight clustering of both homozygous allele classes, while robust calling of the artificially admixed DNA confirmed the KASP marker to be co-dominant (Fig. 4). The marker was subsequently used to genotype a collection of 47 elite UK winter wheat varieties that post-dated the AM panel, released on the UK AHDB Recommended List (https://cereals.ahdb.org.uk/varieties/ahdb-recommended-lists.aspx) between 2009 and 2017 (Supplementary Table 7). Of these, 39 were found to be A:A homozygous alleles, while 8 were G:G (predictive of ToxB insensitivity and sensitivity, respectively). Phenotyping these 47 recent varieties for ToxB sensitivity found the KASP allele call to correctly predict phenotype in all cases (Supplementary Table 7).

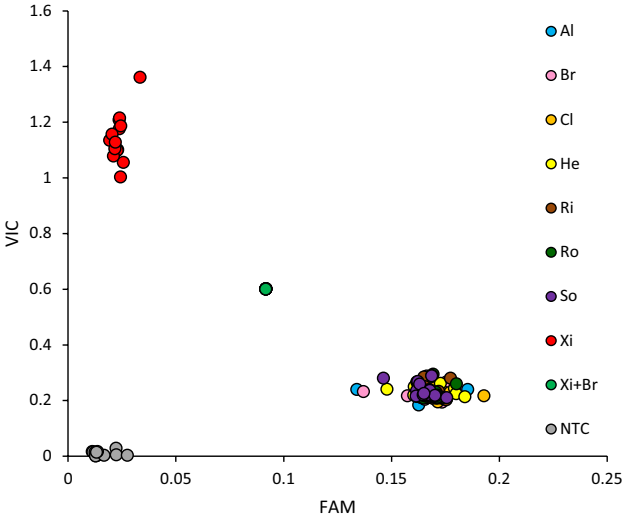


Fig. 4 Allele calls for the *Tsc2*-linked KASP marker developed for SNP BS00072620_51. Fluorescence intensity detected for the FAM (representing nucleotide call G:G, associated with ToxB sensitivity, score ≥ 1) and VIC (A:A, ToxB insensitive, score < 1) fluorophores are shown on the x-axis and y-axis, respectively. Genotyping was undertaken using the DNA extracted from the 'NIAB Elite MAGIC' founders. Al = 'Alchemy', Br = 'Brompton', Cl = 'Claire', He = 'Hereward', Ri = 'Rialto', Ro = 'Robigus', So = 'Soissons', Xi = 'Xi19', Xi + Br = marker co-dominance tested by inclusion of a 50:50 mix of 'Xi19' and 'Brompton' DNA. NTC = no template negative control

Wheat pedigree analysis of ToxB sensitivity

Overlaying ToxB sensitivity scores from the AM panel and post-AM panel onto an up-to-date wheat pedigree (Fradgley et al. 2019) allowed the likely transmission of the ToxB-sensitive allele(s) to be tracked (Fig. 5; Supplementary Table 2). The oldest ToxB-sensitive accession identified was the spring variety 'Thatcher' (released in the USA in 1934), followed by 'Aronde' (France, 1962) and 'Maris Dove' (UK, 1971). Analysis of the pedigree found three varieties to act as 'hubs' for frequent transmission of ToxB sensitivity: (1) 'Xi19' (UK, year of release = 2000, sensitivity = 2.9) had six sensitive descendants ('Cocoon', 'Cubanita', 'Curlew', 'Gallant', 'KWS Curlew' and 'Panorama'), collectively representing 15% of the sensitive European varieties identified. (2) 'Cordiale' (UK, 2001, sensitivity = 1.2) had 12 sensitive descendants, representing around a quarter of ToxB-sensitive accessions. (3) The winter wheat variety 'Aardvark' (UK, 1997), parent or grandparent of 12 sensitive varieties, representing 27% of all ToxB-sensitive lines identified. Despite its prominence in the pedigree of ToxB-sensitive varieties, 'Aardvark' was found to be insensitive. All three 'hub' varieties shared features in common: all were formally found to have ToxB-insensitive parents, and all had 'Cadenza' or a sib of 'Cadenza' listed as one of their parents (Fig. 5). It should be noted that 'Maris Dove' and its progeny 'Axona' are sensitive, and 'Axona' is a

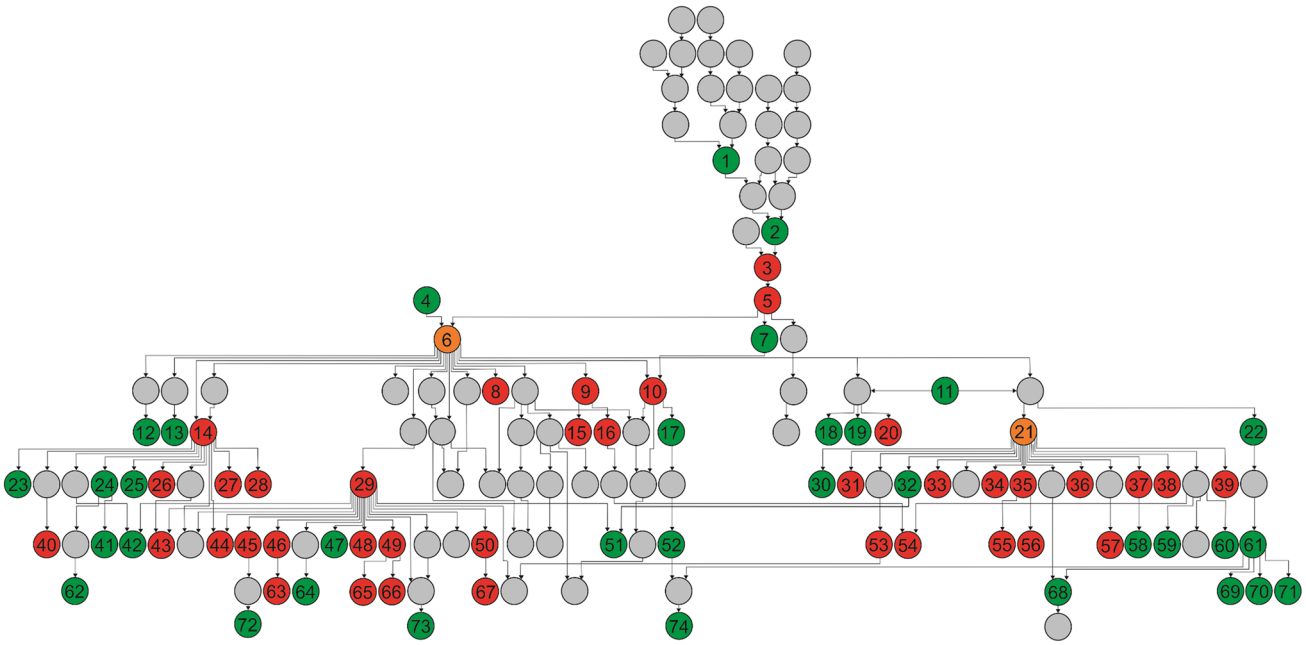


Fig. 5 ToxB sensitivity in the context of wheat pedigree information. ToxB-sensitive (score ≥ 1) and insensitive (score < 1) lines are highlighted in red and green, respectively. Lines or steps in the parentage for which no ToxB phenotype data were available are indicated in grey. The varieties 'Cadenza' (node 6) (and its sibs) and 'Aardvark' (node 21) are highlighted in orange and are hypothesised here to have segregated for allelic status at the *Tsc2* locus during the early period they were used for breeding, with insensitive alleles having subsequently been fixed in commercially released germplasm for both lines. (1) Garnet, (2) Koga II, (3) Maris Dove, (4) Tonic, (5) Axona, (6) Cadenza, (7) Shiraz, (8) Warlock 24, (9) Scorpion 25, (10) Raffles, (11) Lynx, (12) Phlebas, (13) Convoy (14) Xi19, (15) Limerick, (16) Duxford, (17) Cantata (sib), (18) Brando, (19) Ashanti, (20) Arriva, (21) Aardvark, (22) Aardvark (sib), (23) Velocity, (24) Bantham, (25) Rocky, (26) Gallant, (27) Cocoon, (28) Walpole, (29) Cordiale, (30) Galtic, (31) Cadogan, (32) Gulliver, (33) Aarden, (34) Bowindo, (35) Battalion, (36) Scandia, (37) Hyperion, (38) Marksman, (39) Choice, (40) Panorama, (41) KWS Barrel, (42) KWS Podium, (43) Cubanita, (44) KWS Curlew, (45) Grafton, (46) KWS Quartz, (47) Acrobat, (48) Orbit, (49) KWS Sterling, (50) KWS Horizon, (51) Crusoe, (52) Ambrosia, (53) Dover, (54) Moulton, (55) Bennington, (56) RGT Illustrious, (57) Buzzer, (58) Gravitas, (59) Freiston, (60) Dunston, (61) Oakley, (62) LG Mowtown, (63) KWS Zyatt, (64) Costello, (65) KWS Siskin, (66) KWS Silverstone, (67) KWS Lilli, (68) KWS Santiago, (69) Reflection, (70) KWS Gator, (71) KWS Kielder, (72) KWS Trinity, (73) Ranger, (74) RGT Conversion. A high-resolution image in which the variety names for each node in the pedigree are included directly in the figure is available for download as a Supplementary Figure 2

parent of 'Cadenza'. Together, this indicated the most plausible transmission route of ToxB sensitivity to the hub varieties 'Xi19' and 'Cordiale' is via 'Cadenza' breeding lines/sibs still segregating at the *Tsc2* locus. As the remaining hub variety 'Aardvark' has 'Cadenza sib' as one of its parents, and where data were available, none of the remaining parents for the 12 ToxB-sensitive 'Aardvark' descendants were found to possess sensitivity, we speculate that 'Aardvark' germplasm used for breeding was also segregating at *Tsc2*. The assumption of segregation at *Tsc2* in 'Cadenza' and 'Aardvark' germplasm used during their early use in breeding results in all but eight of the 46 ToxB-sensitive varieties being linked in a single pedigree, within which ToxB sensitivity arose from 'Maris Dove'.

Discussion

Sensitivity of north-western European wheat to ToxB

Despite the cloning of *ToxB* over 17 years ago (Martinez et al. 2001), relatively few screens for ToxB sensitivity in wheat have been published. Surveys of Canadian wheat varieties found sensitivity to be present in around 30% of wheat varieties (Tran et al. 2017: 24 of 100 cultivars. Lamari et al. 2005a: 30 of 86 cultivars), while in a recent screen of 122 Australian varieties just 4% were ToxB sensitive (See et al. 2019). Here, screening an AM panel of 480 varieties found ToxB sensitivity to be uncommon in the European germplasm screened, with just 7% possessing a sensitivity score ≥ 1. It appears that ToxB sensitivity was introduced via the variety 'Thatcher', released in the USA in 1931. 'Thatcher' was used 41 times as a parent in

the pedigree, and was a parent of the second oldest sensitive variety in the AM panel, 'Aronde' (France, 1962). 'Thatcher' has been previously identified as the possible source of ToxB sensitivity in Canadian wheat (Lamari et al. 2005a; Tran et al. 2017), highlighting its prominent role in wheat pedigrees across the world.

Analysis of ToxB sensitivity within the context of the wheat pedigree found three varieties to represent hubs for frequent transmission of ToxB sensitivity: 'Xi19' ('Cadenza' × 'Rialto'), 'Cordiale' ('Cadenza' × 'Reaper') and 'Aardvark' (['Cadenza sib' × 'Lynx'] × 'Lynx'). However, a good indication was found that the germplasm for 'Cadenza' and its sibs used during early breeding activities segregated for alleles at *Tsc2*. This was based on the following evidence: (1) the inability to trace the source of sensitivity in the immediate parents of all three hub varieties, (2) the observation that all hub varieties contained 'Cadenza' in their immediate parentage, (3) the ToxB-sensitive variety 'Arriva' represents a 'Lynx' (insensitive) × 'Cadenza' cross, and (4) 'Cadenza' has the sensitive variety 'Axona' as one of its parents. A similar assumption was made for the key hub variety 'Aardvark', whose parentage includes a sib of 'Cadenza' × 'Lynx' (insensitive), and which ToxB phenotyping found to be insensitive, despite being a parent of 12 varieties for which no other source of sensitivity was found. On the assumption of segregation at the *Tsc2* locus for the 'Cadenza' and 'Aardvark' germplasm used during breeding, all but eight ToxB-sensitive lines could be traced via the pedigree to the spring wheat variety 'Maris Dove', released in the UK in 1971. Of the eight ToxB-sensitive lines not accounted for in the pedigree, one was a coded breeder's line with unknown parentage, three contained coded breeders lines in their pedigrees preventing further investigation, one contained parents with unknown sensitivity ('Emerald'. Parents: 'Tadipor' × 'Macao'), one had insensitive parents ('Cyber'. Parents: 'Lynx' × ('Talon' × 'Beaver'), one was the sensitive landrace 'Thatcher', and the last ('Ardone') had 'Thatcher' in its immediate parentage.

It is possible that the low frequency of *Tsc2*-sensitive alleles in our European germplasm may reflect low occurrence of *Ptr* isolates carrying the *ToxB* gene. Indeed, *ToxB* was not present in any of 42 UK *Ptr* isolates we have recently surveyed (J. Cockram and J. Turner unpublished), indicating that sensitivity to this effector is not currently an issue in UK agricultural environments. Surveys of *Ptr* isolates from several countries have shown that presence or absence of the *ToxB* gene varies by region. For example, *ToxB* has been shown to be absent in *Ptr* from Australia (from a screen of 119 isolates; Antoni et al. 2010), New Zealand (12 isolates; Weith 2015) and Latvia, Lithuania and Romania (223 isolates; Abdullah et al. 2017), but present in Canada (Lamari et al. 2005a), Algeria (Benslimane 2018; Lamari et al. 1995; 2003), Azerbaijan, Turkey (Lamari et al. 2003), Syria, Turkey (Lamari et al. 2003; 2005b) and the USA (Abdullah 2017; Ali et al. 1999). Given that modern wheat breeding programmes often incorporate germplasm from around the world, knowledge of varietal differences in ToxB sensitivity will inform wheat breeding in regions where *Ptr* isolates carry the *ToxB* gene, or in regions predicted to be prone to incursions of *ToxB* carrying *Ptr* isolates.

Although neither of the two tetraploid wheat accessions screened here were sensitive to ToxB, the *Tsc2*-ToxB interaction has previously been reported in *T. durum*. This was first shown in the International Triticeae Mapping Initiative (ITMI) hexaploid wheat bi-parental mapping population, in which ToxB sensitivity is conferred by the synthetic hexaploid wheat line 'W7984' (Friesen and Faris 2004). As 'W7984' was developed from a cross between *T. durum* variety 'Altar 84' (the AB sub-genome donor) and *Aegilops tauschii* accession 'CI 18' (D sub-genome donor), the ToxB sensitivity on chromosome 2BS identified in the ITMI population must have originated from *T. durum*. More recently, QTL analysis of ToxB sensitivity and *Ptr* race 5 isolates in an 'Altar' × 'Langdon' durum population has demonstrated the role of ToxB sensitivity conferred by the *Tsc2* locus with tan spot susceptibility in *T. durum* (Virdi et al. 2016).

Genetic control of ToxB sensitivity in wheat.

QTL mapping using the MAGIC and AM panels found ToxB sensitivity to be predominantly controlled by *Tsc2*. The MAGIC proximal marker originated from a wheat gene model that also contained GWAS hits 1–5 and 10. However, analysis of haplotypes around the *Tsc2* region in the AM panel unambiguously positioned *Tsc2* distal to gene model TraesCS2B01G046400 (based on SNPs Kukri_c148_1512, Kukri_c148_1346, representing GWAS hits 7 and 8, respectively) based on recombination events in 'Orbit' and 'Cadogan'. The MAGIC marker immediately distal to the *Tsc2* QTL peak corresponded to AM hit 12 (Kukri_c63748_1453), with analysis of haplotypes in the AM panel positioning this marker distal to *Tsc2,* based on recombination events in four varieties: 'Selkirk', 'Elka', 'Thatcher' and 'Aronde' (Supplementary Tables 2 and 4). Therefore, the genomic regions defined by GWAS in the AM panel and genetic analyses in the MAGIC population combined to localise *Tsc2* to a 1921-kb region containing 104 gene models, based on the wheat cv. 'Chinese Spring' reference genome.

Within the *Tsc2* region, SNP BS00072620_51 was found to be the most significant marker in both of the genetic mapping populations investigated. Furthermore, it was shown in the AM panel ($n = 480$) to be diagnostic for ToxB sensitivity, with A:A and G:G alleles perfectly predicting insensitivity (ToxB score < 1) or sensitivity (score ≥ 1), respectively.

Similarly, SNP BS00072620_51 was also diagnostic for ToxB sensitivity in the 47 varieties investigated that post-dated the AM panel. As ToxB sensitivity has been shown to be correlated with susceptibility to tan spot (Friesen and Faris 2004; Abeysekara et al. 2010; Singh et al. 2010), the co-dominant KASP marker developed here which is diagnostic for allelic state at *Tsc2* represents a useful tool for selecting against ToxB sensitivity in wheat breeding programmes in regions where ToxB-containing *Ptr* is prevalent. The incursion of isolates carrying ToxB into new areas would make marker-assisted selection against ToxB sensitivity a priority target for guarding against future changes in pathogen prevalence and virulence.

The putative minor QTL identified in the MAGIC population on the long arm of chromosome 2A differs from previously identified QTL controlling sensitivity to *Ptr* race 5 isolates located on chromosomes 2BS (*Tsr6*, Singh et al. 2010), 2AS and 2BL (Friesen and Faris 2004). Due to the low percentage of the phenotypic variation accounted for, and its identification only in the CIM–cov10 analyses, the status of this QTL is currently putative. While further investigation of this minor QTL would initially only have been thought to be of interest in determining the genetic pathways controlling the fine control of wheat ToxB sensitivity, it is interesting to note that this QTL co-locates with a robust QTL in the same MAGIC population conferring both adult plant and seedling resistance to the related necrotrophic pathogen *P. nodorum* (*QSnb.niab-2A.3* at ~ 140 cM) (Lin et al. 2019). Similarly, a minor QTL controlling sensitivity to the *P. nodorum* effector Tox3 (Downie et al. 2018) is reported to collocate with another chromosome 2A *P. nodorum* field resistance QTL, *Qsnb.niab.2A.4* (Lin et al. 2019). Collectively, these results indicate that minor necrotrophic fungal pathogen effector sensitivity QTL may also be relevant to field resistance to this class of wheat necrotrophic fungal pathogens, as has been shown for the major *P. nodorum* effector sensitivity loci *Tsn1, Snn1* and *Snn3-B1* in wheat (e.g. Friesen et al. 2009; Phan et al. 2016; Ruud et al. 2017).

Disease resistance breeding and selection against effector sensitivity loci

The eight *Ptr* races identified are differentiated by their expression of one or a combination of the three known effectors (Lamari et al. 2003). Thus, in the context of effector sensitivity, *Ptr*–wheat interactions will depend on the allelic states at the respective loci in both the host and pathogen. It should be noted that it is likely that additional effectors remain to be discovered, and that other factors also contribute to host resistance to *Ptr* infection. Furthermore, work in another necrotrophic fungal pathogen of wheat, *P. nodorum*, has shown that the expression of one effector can suppress that of another (Phan et al. 2016). Therefore, a more complex genetic control may result from an otherwise seemingly simple underlying system. For example, while we find Chinese Spring to be insensitive to ToxB, this variety does show varying resistance/susceptibility to different *Ptr* isolates. Indeed, Tadesse et al (2006) used a *Ptr* isolate found to be virulent on Chinese Spring to screen germplasm generated by crossing 100 synthetic hexaploid wheat accessions to Chinese Spring nullisomic lines, identifying several resistance loci on chromosome 3D. Ultimately, the best strategies for increasing resistance to the widest diversity of *Ptr* strains will likely involve combined approaches, in which removal of effector sensitivity loci is complemented by genetic mapping of tan spot resistance QTL and use of linked markers for marker-assisted selection of favourable alleles.

Future prospects towards cloning Tsc2

The current consensus is that ToxB sensitivity is predominantly conferred by a single locus, *Tsc2* (Friesen and Faris 2004; Abeysekara et al. 2010). The wheat reference genome has been constructed in the variety 'Chinese Spring', which is insensitive to ToxB. The only other *Ptr* effector sensitivity locus that has been cloned is *Tsn1*, at which recessive alleles confer insensitivity to the effector ToxA. The protein kinase/NBS-LRR gene underlying *Tsn1* is deleted in almost all ToxA-insensitive varieties (Faris et al. 2010), supporting the hypothesis that the gene underlying *Tsc2* may also be absent in ToxB-insensitive varieties, including 'Chinese Spring'. To date, a small number of hexaploid wheat varieties have been resequenced, including 'Claire', 'Paragon', 'Robigus' (available at https://wheatis.tgac.ac.uk/grassroots-portal/blast). However, we found none of these varieties to be sensitive to ToxB. Here we investigated the genome sequence of the synthetic hexaploid wheat accession 'W7984' (Chapman et al. 2015), previously shown to carry a ToxB-sensitive allele at the *Tsc2* locus (Friesen and Faris 2004). While the 'W7984' contigs have low (55%) genome coverage and have short length (contig N50 6.7 kb), we nevertheless thought it useful to undertake an analysis of the potential gene space in a ToxB-sensitive line. This analysis allowed the identification of additional candidate genes within the ToxB-sensitive line 'W7984' super-contig that were not present in ToxB-insensitive variety 'Chinese Spring', providing potential leads for future analysis. We did not further our analysis of the 'W7984' genome sequence due to its fragmented nature, and because a number of projects are currently resequencing additional hexaploid bread wheat varieties. These include the 10 + Wheat Genomes Project (https://www.10wheatgenomes.com/), as well as an ongoing project to resequence all eight parents of the 'NIAB Elite MAGIC' population (https://gtr.ukri.org/projects?ref=BB%2FP010741%2F1). As MAGIC founder 'Xi19' is ToxB sensitive, we hope that

the future completion of the 'Xi19' genome assembly will aid identification of the genetic variant(s) underlying *Tsc2*.

Acknowledgements This work was undertaken within the framework of the 2nd call ERA-NET for Coordinating Plant Sciences, within the 'Efectawheat' project, funded by British Biological Sciences Research Council (BBSRC) grant BB/N00518X/1. Additional funding was via Grains and Research Development Corporation (GRDC) grant CUR00023. LV's time was funded by BBSRC grant BB/P010768/2. We thank Bostjan Kobe (University of Queensland) and Simon Williams (Australian National University) for the SHuffle strain. We thank NIAB colleagues Nick Fradgley for help with visualisation of the pedigree information and Keith Gardner for advice with genetic analyses and review of the manuscript.

Author Contribution statement JC, CSM and RPO gained funding and conceived the project. EI and PS produced the ToxB protein. BC, RPO, CSM and CM-B undertook the ToxB infiltration experiments. JC, RD, LV, CCM and LP undertook bioinformatic analyses. BC designed and performed all other research and analysis. JC, RPO and CSM provided supervision and project resources. BC and JC wrote the manuscript. All authors edited and approved the manuscript.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Abdullah S (2017) Characterization of *Pyrenophora tritici-repentis* in wheat and rye to study tan spot susceptibility and insights into its relationship with stem rust resistance. Theses and Dissertations. 1183

Abdullah S, Sehgal SK, Ali S, Liatukas Z, Ittu M, Kaur N (2017) Characterization of *Pyrenophora tritici-repentis* (tan spot of wheat) races in Baltic states and Romania. Plant Pathol J 33:133–139

Abeysekara NS, Friesen TL, Liu Z, McClean PE, Faris JD (2010) Marker development and saturation mapping of the Tan Spot Ptr ToxB sensitivity locus *Tsc2* in hexaploid wheat. Plant Genome 3:179–189

Ali S, Francl LJ, Dewolf ED (1999) First report of *Pyrenophora tritici-repentis* race 5 from North America. Plant Dis 83:591

Allen AM, Winfield MO, Burridge AJ, Downie RC, Benbow H, Barker GLA et al (2017) Characterization of a wheat breeders' array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). Plant Biotechnol J 15:390–401

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Annone J (1998) Tan spot of wheat in Argentina: importance of disease management strategies. In: Duveiller E, Dubin H, Reeves J, McNab A (eds) Mexico. D.F, CIMMYT, pp 339–345

Antoni EA, Rybak K, Tucker MP, Hane JK, Solomon PS, Drenth A, Shankar M, Oliver RP (2010) Ubiquity of ToxA and absence of ToxB in Australian populations of *Pyrenophora tritici-repentis*. Australian Plant Pathol 39:63–68

Ballance GM, Bernier CC, Lamari L (1989) Purification and characterization of a host-selective necrosis toxin from *Pyrenophora tritici-repentis*. Physiol Mol Plant Pathol 35:203–213

Benslimane H (2018) Virulence phenotyping and molecular characterisation of a new virulence type of *Pyrenophora tritici-repentis* the causal agent of Tan Spot. Plant Pathol J 34:139–142

Broman KW, Wu H, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890

Chapman JA, Mascher M, Buluç A, Barry K, Georganas E, Session A et al (2015) A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. BMC Genome Biol 16:26

Ciuffetti LM, Tuori RP, Gaventa JM (1997) A single gene encodes a selective toxin causal to the development of tan spot of wheat. Plant Cell 9:135–144

Ciuffetti LM, Manning VA, Pandelova I, Betts M, Martinez JP (2010) Host-selective toxins, PtrToxA and PtrToxB, as necrotrophic effectors in *Pyrenophora tritici-repentis*—wheat interaction. New Phytol 187:911–919

Ciuffetti LM, Manning VA, Pandelova I, Faris JD, Friesen TL, Strelkov SE, Weber GL, Goodwin SB, Wolpert TJ, Figueroa M (2014) Pyrenophora tritici-repentis: a plant pathogenic fungus with global impact. Genomics of plant-associated fungi: monocot pathogens. Springer, Berlin, pp 1–39

Cockram J, Scuderi A, Barber T, Furuki E, Gardner KA, Gosman N, et al. (2015) Fine-mapping the wheat Snn1 locus conferring sensitivity to the Parastagonospora nodorum necrotrophic effector SnTox1 using an eight founder multi-parent advanced generation intercross population. G3 Genes Genomes Genet 5:2257–2266.

De Wolf ED, Efferetz RJ, Ali S, Feancl LJ (1998) Vistas of tan spot research. Can J Plant Pathol 20:349–370

Downie RC, Bouvet L, Furuki E, Gosman N, Gardner KA, Mackay IJ, Mantello CC, Mellers G, Phan HTT, Rose GA, Tan K-C, Oliver RP, Cockram J (2018) Assessing European wheat sensitivities to *Parastagonospora nodorum* necrotrophic effectors and fine-mapping of the *Snn3-B1* locus conferring sensitivity to the effector SnTox3. Front Plant Sci 9:881

Effertz RJ, Meinhardt SW, Anderson JA, Jordahl JG, Francl LJ (2002) Identification of chlorosis-inducing toxin from *Pyrenophora tritici-repentis* and the chromosomal location of an insensitivity locus in wheat. Phytopathology 92:527–533

Faris JD, Anderson JA, Francl LJ, Jordahl JG (1996) Chromosomal location of a gene conditioning insensitivity in wheat to a necrosis-inducing culture filtrate from *Pyrenophora triticirepentis*. Phytopathology 86:459–63

Faris JD, Zhang Z, Lu H, Reddy L, Clouter S, Fellers JP et al (2010) A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens. Proc Natl Acad Sci USA 107:13544–13549

Fradgley N, Gardner KA, Cockram J, Elderfield J, Hickey JA, Howell P et al (2019) A large-scale pedigree resource of wheat reveals evidence for adaptation and selection by breeders. PLoS Biol 17:e3000071

Friesen TL, Faris JD (2004) Molecular mapping of resistance to *Pyrenophora tritici-repentis* race 5 and sensitivity to Ptr ToxB in wheat. Theor Appl Genet 109:464–471

Friesen TL, Ali S, Klein K, Rasmussen JB (2005) Population genetic analysis of a global collection of *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. Phytopathology 95:1144–1150

Friesen TL, Stukenbrock EH, Liu Z, Meinhardt S, Ling H, Faris JD, Rasmussen JB, Solomon PS, McDonald BA, Oliver RP (2006) Emergence of a new disease as a result of interspecific virulence gene transfer. Nat Genet 38:953–956

Friesen TL, Chu CG, Liu ZH, Xu SS, Halley S, Faris JD (2009) Host-selective toxins produced by *Stagonospora nodorum* confer disease susceptibility in adult wheat plants under field conditions. Theor Appl Genet 118:1489–1497

Fulton TM, Chunwongse T, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13:207–209

Gardner KA, Wittern LM, Mackay IJ (2016) A highly recombined, high-density, eight founder wheat MAGIC map reveals extensive segregation distortion and genomic locations of introgression segments. Plant Biotechnol J 14:1406–1417

Gremme G, Brendel V, Sparks ME, Kurtz S (2005) Engineering a software tool for gene structure prediction in higher organisms. Inf Softw Technol 47:965–978

Huang BE, George AW (2011) R/mpMap: a computational platform for the genetic analysis of multiparent recombinant inbred lines. Bioinformatics 27:727–729

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ et al (2008) Efficient control of population structure in model organism association mapping. Genetics 178:1709–1723

Kim YM, Bouras N, Kav NN, Strelkov SE (2010) Inhibition of photosynthesis and modification of the wheat leaf proteome by Ptr ToxB: a host-specific toxin from the fungal pathogen *Pyrenophora tritici-repentis*. Proteomics 10:2911–2926

Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissod L et al (2017) Uncovering hidden variation in polyploid wheat. Proc Natl Acad Sci USA 114:E913–E921

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645

Lamari L, Sayoud R, Boulif M, Bernier CC (1995) Identification of a new race in *Pyrenophora tritici-repentis*: implications for the current pathotype classification system. Can J Plant Pathol 17:312–318

Lamari L, Strelkov S, Yahyaoui J, Orabi J, Smith RB (2003) The identification of two new races of *Pyrenophora tritici-repentis* from the host center of diversity confirms a one-to-one relationship to tan spot of wheat. Phytopathology 93:391–396

Lamari L, McCallum BD, Depauw RM (2005a) Forensic pathology of Canadian bread wheat: the case of tan spot. Phytopathology 95:144–152

Lamari L, Strelkov SE, Yahyaoui A, Amedov M, Saidov M, Djunusova M, Koichibayev M (2005b) Virulence of *Pyrenophora tritici-repentis* in the countries of the Silk Road. Can J Plant Pathol 27:383–388

Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100

Lin M, Corsi B, Ficke A, Cockram J, Lillemo M (2019) Genetic mapping using a wheat multi-founder population reveals a locus on chromosome 2A controlling resistance to both leaf and glume blotch caused by the necrotrophic fungal pathogen *Parastagonospora nodorum*. Theor Appl Genet (submitted)

Lipka AE, Tian F, Wang Q, Peiffer J, Li M et al (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28:2397–2399

Lu S, Faris JD, Sherwood R, Friesen TL, Edwards MC (2014) A dimeric PR-1-type pathogenesis-related protein interacts with ToxA and potentially mediates ToxA-induced necrosis in sensitive wheat. Mol Plant Pathol 15:650–663

Mackay IJ, Bansept-Basler P, Barber T, Bentley AR, Cockram J, Gosman N, et al. (2014) An eight-parent multiparent advanced generation inter-cross population for winter sown wheat: creation, properties, and validation. G3 Genes Genomes Genet 4:1603–1610.

Manning VA, Hardison LK, Ciuffetti LM (2007) Ptr ToxA interacts with a chloroplast-localized protein. Mol Plant Microbe Interact 20:168–177

Manning VA, Hamilton SM, Karplus PA, Ciuffetti LM (2008) The Arg-Gly-Asp-containing, solvent-exposed loop of Ptr ToxA is required for internalization. Molec Plant Microbe Interact 21:315–325

Martinez JP, Ottum SA, Ali S, Franci LJ, Ciuffetti LM (2001) Characterization of the ToxB gene from *Pyrenophora tritici-repentis*. Mol Plant Microbe Interact 14:675–677

Martinez JP, Oesch NW, Ciuffetti LM (2004) Characterization of the multiple-copy host-selective toxin gene, ToxB, in pathogenic and nonpathogenic isolates of *Pyrenophora tritici-repentis*. Mol Plant Microbe Interact 17:467–474

Meinhardt SW, Cheng W, Kwon CY, Donohue CM, Rasmussen JB (2002) Role of the Arginyl-Glycyl-Aspartic motif in the action of Ptr ToxA produced by *Pyrenophora tritici-repentis*. Plant Physiol 130:1545–1551

Moffat CS, See PT, Oliver RP (2014) Generation of a ToxA knockout of the wheat tan spot pathogen *Pyrenophora tritic-repentis*. Molec Plant Breed 15:918–926

Phan HT, Rybak K, Furuki E, Breen S, Solomon PS, Oliver RP, Tan KC (2016) Differential effector gene expression underpins epistasis in a plant fungal disease. Plant J 87:343–354

R Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ramirez-Gonzalez RH, Uauy C, Caccamo M (2015) PolyMarker: a fast polyploid primer design pipeline. Bioinformatics 31:2038–2039

Ruud AK, Windju S, Belova T, Friesen TL, Lillemo M (2017) Mapping of SnTox3-*Snn3* as a major determinant of field susceptibility to Septoria nodorum leaf blotch in the SHA3/CBRD × Naxos population. Theor Appl Genet 130:1361–1374

Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A (2019) The global burden of pests on major food crops. Nat Ecol Evol 3:430–439

See PT, Marathamuthu EM, Iagallo EM, Oliver RP, Moffat CS (2018) Evaluating the importance of the tan spot ToxA–*Tsn1* interaction in Australian wheat varieties. Plant Pathol 67:1066–1075

See PT, Iagallo EM, Oliver RP, Moffat CS (2019) Heterologous expression of the *Pyrenophora tritici-repentis* effector proteins ToxA and ToxB, and the prevalence of effector sensitivity in Australian cereal crops. Front Microbiol 10:182

Shaw PD, Kennedy J, Graham M, Milne I, Marshall DF (2014) Helium: visualization of large scale plant pedigrees. BMC Bioinform 15:259

Singh PK, Mergoum M, Adhikari TB, Shah T, Ghavami F, Kianian SF (2010) Genetic and molecular analysis of wheat tan spot resistance effective against *Pyrenophora tritici-repentis* races 2 and 5. Mol Breed 25:369–379

Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC (1985) Measurement of protein using bicinchoninic acid. Anal Biochem 150:76–85

Stadlmeier M, Harlt L, Mohler V (2018) Usefulness of a multiparent advanced generation intercross population with a greatly reduced mating design for genetic studies in winter wheat. Front Plant Sci 9:1825

Stanke M, Morgenstern B (2005) AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res 33:W465–W467

Strelkov SE, Lamari L (2003) Host-parasite interactions in tan spot (*Pyrenophora tritici-repentis*) of wheat. Can J Plant Pathol 25:339–349

Strelkov SE, Lamari L, Ballance GM (1998) Induced chlorophyll degradation by a chlorosis toxin from *Pyrenophora tritici-repentis*. Can J Plant Pathol 20:428–435

Strelkov SE, Lamari L, Sayoud R, Smith RB (2002) Comparative virulence of chlorosis-inducing races of *Pyrenophora tritici-repentis*. Can J Plant Pathol 24:29–35

Tadesse W, Hsam SLK, Wenzel G, Zeller FJ (2006) Identification and monosomic analysis of tan spot resistance genes in synthetic wheat lines (Triticum turgidum L. × Aegilops tauschii Coss.). Crop Sci 46:1212–1217

Tai YS, Bragg J, Meinhardt SW (2007) Functional characterization of ToxA and molecular identification of its intracellular targeting protein in wheat. Am J Plant Physiol 2:76–89

Tan KC, Oliver RP, Solomon PS, Moffat CS (2010) Proteinaceous necrotrophic effectors in fungal virulence. Funct Plant Biol 37:907–912

The International Wheat Genome Sequencing Consortium (IWGSC), Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J Pozniak C, et al. (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361:eaar7191

Tomas A, Gene GH, Reeck GR, Bockus WW, Leach JE (1990) Purification of a cultivar-specific toxin from *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. Mol Plant Microbe Interact 3:221–224

Tran VA, Aboukhaddour R, Strelkov IS, Bouras N, Spaner D, Strelkov SE (2017) The sensitivity of Canadian wheat genotypes to the necrotrophic effectors produced by *Pyrenophora tritici-repentis*. Can J Plant Pathol 39:149–162

Venturini L, Caim C, Kaithakottil GG, Mapleson DL, Swarbreck D (2018) Leveraging multiple transcriptome assembly methods for improved gene structure annotation. Gigascience 7:giy093

Virdi SK, Overlander ME, Zhang Z, Xu SS, Friesen TL, Faris JD (2016) New insights into the roles of host gene-necrotrophic effector interactions in governing susceptibility of durum wheat to tan spot and septoria nodorum blotch. G3 Genes Genom Genet 6:4139–4150

VSN International (2015) Genstat for windows. VSN International, Hemel Hempstead, 18th edn. UK. webpage: www.genstat.co.uk.

Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE et al (2014) Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. Plant Biotechnol J 12:787–796

Weith K (2015) *Pyrenophora tritici-repentis* the causal agent of tan spot: characterisation of New Zeland populations. MSc thesis. Lincoln University, United Kingdom

Zadoks Z, Chang TT, Konzak CF (1974) A decimal code for the growth stages of cereals. Weed Res 14:415–421

Zhang Z, Ersoz E, Lai C-Q, Todhunter RJ, Tiwari HK, Gore MA et al (2010) Mixed linear model approach adapted for genome-wide association studies. Nat Genet 42:355–360